

Subcritical multiplicative chaos for regularized counting statistics from random matrix theory

Gaultier Lambert*, Dmitry Ostrovsky† and Nick Simm‡

December 8, 2016

Abstract

For an $N \times N$ random unitary matrix U_N , we consider the random field defined by counting the number of eigenvalues of U_N in a mesoscopic arc of the unit circle, regularized at an N -dependent scale $\epsilon_N > 0$. We prove that the renormalized exponential of this field converges as $N \rightarrow \infty$ to a Gaussian multiplicative chaos measure in the whole subcritical phase. In addition, we show that the moments of the total mass converge to a Selberg-like integral and by taking a further limit as the size of the arc diverges, we establish part of the conjectures in [54]. By an analogous construction, we prove that the multiplicative chaos measure coming from the sine process has the same distribution, which strongly suggests that this limiting object should be universal. The proofs are based on the asymptotic analysis of certain Toeplitz or Fredholm determinants using the Borodin-Okounkov formula or a Riemann-Hilbert problem for integrable operators. Our approach to the L^1 -phase is based on a generalization of the construction in Berestycki [4] to random fields which are only *asymptotically* Gaussian. In particular, our method could have applications to other random fields coming from either random matrix theory or a different context.

1 Introduction

1.1 Background and motivation

The study of Gaussian fields with logarithmic correlations has seen many recent developments. One of those concerns a non-trivial relation with the eigenvalues of random matrices, which first appeared explicitly in work of Hughes, Keating and O’Connell [34]. They studied the characteristic polynomial of $N \times N$ unitary matrices U_N drawn at random from the unitary group, with Haar measure, also known as the Circular Unitary Ensemble (CUE). One of their key results was to prove that $Z_N(\theta) = \sqrt{2} \log |\det(e^{i\theta} - U_N)|$ converges in law as $N \rightarrow \infty$ to a log-correlated Gaussian field which corresponds to the restriction of the two-dimensional Gaussian Free Field on the unit circle. This field can be represented as the Fourier series:

$$Z(\theta) = \frac{1}{\sqrt{2}} \operatorname{Re} \left\{ \sum_{k=1}^{\infty} \frac{e^{-ik\theta}}{\sqrt{k}} Z_k \right\}, \quad (1.1)$$

where Z_k are i.i.d. standard complex Gaussian variables. In particular, almost surely, this series does not converge in $L^2(\mathbb{T})$, and $Z(\theta)$ only exists as a random distribution, or a generalized random function with correlation kernel $\mathbb{E}(Z(\theta)Z(\theta')) = -\log |e^{i\theta} - e^{i\theta'}|$.

*Department of Mathematics, Universität Zürich, gaultier.lambert@math.uzh.ch. G. L. gratefully acknowledges the support of the grant KAW 2010.0063 from the Knut and Alice Wallenberg Foundation while at KTH, Royal Institute of Technology.

†dm.ostrov@aya.yale.edu.

‡Mathematics Institute, University of Warwick, n.simm@warwick.ac.uk. N. S. gratefully acknowledges financial support from a Leverhulme Trust Early Career Fellowship ECF2014-304.

More recent developments concern the extreme value statistics of the field $Z_N(\theta)$ and on the related issue of making sense of its exponential in the limit as $N \rightarrow \infty$. The authors of [24, 25] gave a very precise conjecture for the maximum value of $Z_N(\theta)$, which has recently seen significant progress [1, 55, 12, 44]. This conjecture is intimately related to the following¹ concerning the total mass of the field $Z_N(\theta)$.

Conjecture 1.1 (Fyodorov and Keating [25]). *Let $\gamma > 0$ and define*

$$M_N^\gamma := \int_0^{2\pi} e^{\gamma Z_N(\theta)} d\theta \quad (1.2)$$

For any $q \in \mathbb{N}$ such that $\gamma^2 q < 2$, we have

$$\begin{aligned} \lim_{N \rightarrow \infty} N^{-\frac{\gamma^2 q}{2}} \mathbb{E}[(M_N^\gamma)^q] &= C_{\mu, q} \int_{[0, 2\pi]^q} \prod_{1 \leq j < k \leq q} |e^{i\theta_j} - e^{i\theta_k}|^{-\gamma^2} d\theta_1 \dots d\theta_q \\ &= C_{\gamma, q} (2\pi)^q \frac{\Gamma(1 - \gamma^2 q/2)}{\Gamma(1 - \gamma^2/2)^q} \end{aligned} \quad (1.3)$$

where $C_{\gamma, q} = \frac{G(1+\gamma/\sqrt{2})^{2q}}{G(1+\gamma\sqrt{2})^q}$ and $G(z)$ is the Barnes G -function.

The case $q = 2$ was recently solved by Claeys and Krasovsky [13] using sophisticated asymptotics of Toeplitz determinants with merging Fisher-Hartwig singularities. For $q \neq 1, 2$, this conjecture remains an open problem. The integral on the first line of (1.3), known as the Dyson integral, is convergent if and only if $\gamma^2 q < 2$. It is a special case of the Morris integral and belongs to a class of integrals evaluated explicitly by Selberg in 1944, of which the Selberg integral is the most important. See [22] for a comprehensive account of such integrals across mathematics and physics. In particular, it is known that the Dyson and Selberg integrals describe the moments of the total mass of the Bacry-Muzy Gaussian multiplicative chaos measure on the circle and the interval, respectively. Hence the problem of extending these integrals to meromorphic functions of $q \in \mathbb{C}$ in such a way that the resulting function is the Mellin transform of a probability distribution is of fundamental importance as the probability distribution is then naturally conjectured to be that of the total mass. This problem was solved by Fyodorov and Bouchaud [23] for the Dyson integral, heuristically by Fyodorov *et. al.* [28] and independently and rigorously by the second author [50] for the Selberg integral, see also [51] and [53], and [52] for the Morris integral. These analytic extensions are of considerable importance in the theory of log-correlated Gaussian processes [23, 28, 24, 25, 29, 27] as they lead to precise conjectures about the explicit form of the extreme value statistics, by an analogy with the so-called derivative martingale coming from branching processes. For our purposes, the analytic extension of the Selberg integral is directly related to our results for the CUE; see the discussion following theorem 1.3 below.

The conjecture 1.1 is also related to Gaussian Multiplicative Chaos (GMC), a theory devoted to making sense of the exponential of a log-correlated Gaussian field G . This theory was introduced in the 80's to provide a random model to describe the energy dissipation in a turbulent flow and later in connection with 2d quantum gravity, which is a theory developed by Polyakov to make sense of the concept of a random geometry. One of the pioneering mathematical works in the subject is that of Kahane [39, 38] who provided a consistent way to make sense of $e^{\gamma G(x)}$ as a random measure in the case where the correlation kernel of G is said to be *totally positive*. Moreover, he established that this construction yields a non-trivial measure if and only if the parameter $\gamma < \sqrt{2d}$ where d is the dimension. The regime $\gamma < \sqrt{2d}$ is precisely when the mean of the total mass (1.2) is non-zero and it is usually called the subcritical phase. Recently, the theory has seen many significant

¹For convenience we state conjecture 1.1 for $q \in \mathbb{N}$, the full conjecture given by these authors involves an analytic continuation to $q \in \mathbb{C}$ which allowed them to calculate the extreme value statistics of $Z_N(\theta)$

contributions, a small sample of which is presented in section 1.5. This is far from exhaustive and we refer instead to the survey [58] or the lecture notes [5] for a comprehensive introduction. Let us just mention that more general approaches than Kahane's have been developed to define Gaussian multiplicative chaos, including an interesting intrinsic definition given by Shamov in [63], as well as an approach which consists of regularizing at scale $\epsilon > 0$ the field G using mollifiers and taking the limit as $\epsilon \rightarrow 0$ [60, 4]. In this paper, we reutilize this idea to regularize counting statistics coming from random matrices and construct the associated chaos measure. As mentioned earlier, this could have applications for understanding the distribution of the maximum of such statistics.

The connection between random matrices and multiplicative chaos was rigorously put forward in [68]. Webb constructed the multiplicative chaos associated with the field $Z_N(\theta)$ in the L^2 -phase (*i.e.* $\gamma < 1$) and showed that its law agrees with that constructed out of the log-correlated field $Z(\theta)$. In a non-Gaussian setting, the complex multiplicative chaos associated with the Riemann ζ -function was also obtained in [62] by using a coupling with a Gaussian process. One of the purposes of this paper is to extend the construction of the multiplicative chaos to the whole subcritical regime and to apply our method to random matrix statistics. Specifically, we will consider the CUE and the sine process which describes the mesoscopic behavior of unitary invariant ensembles in the bulk [43].

Let $e^{i\theta_1}, \dots, e^{i\theta_N}$ denote the eigenvalues of a Haar distributed unitary matrix with the convention that $\theta_1, \dots, \theta_N \in [0, 2\pi)$. For a parameter $\ell > 0$, consider the random process $u \rightarrow X_N(u)$ which counts the number of eigenangles in an interval centered around u :

$$X_N(u) = \sum_{j=1}^N \chi_u(N^\alpha \theta_j), \quad (1.4)$$

where $\chi_u(x) = \pi \mathbf{1}_{|x-u| \leq \ell/2}$. The parameter α is called the spectral scale of the process (1.4), which can be studied in either the global regime $\alpha = 0$ or the mesoscopic regime $0 < \alpha < 1$. In both regimes and in the limit $N \rightarrow \infty$, the process $X_N(u)$ is typically Gaussian and log-correlated. We will focus on the mesoscopic regime where the geometry of the spectrum is unimportant, although our theory also applies for $\alpha = 0$ with minor changes. Throughout the paper we will work with the following family of regularizations instead of working explicitly with the counting statistics (1.4). Let ϕ be a sufficiently smooth probability density function (or mollifier) and define

$$X_{N,\epsilon}(u) := \sum_{j=1}^N (\chi_u \star \phi_\epsilon)(N^\alpha \theta_j) \quad (1.5)$$

where $\phi_\epsilon(\theta) = \frac{1}{\epsilon} \phi(\frac{\theta}{\epsilon})$. Note that $\epsilon > 0$ controls the scale of regularization for $X_N(u)$. Such smoothed fields were introduced in [54] in the context of counting statistics of the Riemann zeros. The main advantage of smoothing is that it is easier to control the asymptotics of the Laplace transform of (1.5), see remark 1.4. For $\epsilon > 0$ fixed, the centered process $X_{N,\epsilon}(u) - \mathbb{E}(X_{N,\epsilon}(u))$ converges as $N \rightarrow \infty$ to a Gaussian process $X_\epsilon(u)$ whose covariance is given by an $H^{1/2}$ -inner product. In the global regime ($\alpha = 0$), this is a consequence of the strong Szegő limit theorem. In the mesoscopic regime ($0 < \alpha < 1$) this follows from a central limit theorem (CLT) of Soshnikov [66]: For any smooth function h with sufficiently rapid decay at $\pm\infty$, we have

$$\sum_{j=1}^N h(N^\alpha \theta_j) - \mathbb{E} \left(\sum_{j=1}^N h(N^\alpha \theta_j) \right) \rightarrow \mathcal{N} \left(0, \int_{\mathbb{R}} |k| |\hat{h}(k)|^2 dk \right), \quad N \rightarrow \infty, \quad (1.6)$$

where the Fourier transform is normalized as $\hat{h}(k) = \int_{\mathbb{R}} h(x) e^{-2\pi i k x} dx$. Of particular interest for our purposes is that the covariance structure associated with this CLT can be expressed in the

form, for any functions $g, h \in C_0^1(\mathbb{R})$,

$$\begin{aligned} \lim_{N \rightarrow \infty} \text{Cov} \left(\sum_{j=1}^N g(N^\alpha \theta_j), \sum_{j=1}^N h(N^\alpha \theta_j) \right) &= \int_{\mathbb{R}} |k| \hat{g}(k) \hat{h}(-k) dk \\ &= \frac{1}{2\pi^2} \int_{\mathbb{R}^2} g'(x) h'(y) \log \frac{1}{|x-y|} dx dy \end{aligned} \quad (1.7)$$

and is therefore suggestive of a logarithmic covariance structure underlying the mesoscopic statistics of CUE eigenvalues. Applied to the statistic (1.5), one can show that $\text{Var}(X_{N,\epsilon}(u)) \sim \log \frac{1}{\epsilon}$ as $N \rightarrow \infty$ and $\epsilon \rightarrow 0$ and that for $u, v \in \mathbb{R}$ fixed ($u \neq v$), we have

$$\lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \text{Cov}(X_{N,\epsilon}(u), X_{N,\epsilon}(v)) = \log \frac{1}{|u-v|} + \log |\ell^2 - (u-v)^2|^{1/2}. \quad (1.8)$$

In fact, the appearance of the $H^{1/2}$ -Gaussian noise is remarkably universal in the mesoscopic limit of one dimensional random process *with rigidity*. This problem has attracted renewed interest in the last couple of years. For instance, analogous results were obtained for the GUE [26], more general invariant ensembles [11, 43], Wigner matrices [21, 45, 31] as well as to β -ensembles [9, 2] and in connection to zeros of the Riemann zeta function [10, 61].

Based on formula (A.13) in the appendix, the limiting Gaussian process associated to the counting statistics (1.4) is stationary on \mathbb{R} and has spectral density

$$\hat{Q}(\kappa) = \frac{\sin^2(\pi \ell \kappa)}{|\kappa|}. \quad (1.9)$$

This is manifestly an example of $1/f$ -noises or log-correlated field, see [26, 69] and references therein for processes of a similar type. In fact, an explicit computation shows that the corresponding mean-zero Gaussian process, denoted by G throughout this paper, has covariance function:

$$\mathbb{E}[G(u)G(v)] = Q(u-v) = \log \frac{1}{|u-v|} + \log |\ell^2 - (u-v)^2|^{1/2}. \quad (1.10)$$

for any $u, v \in \mathbb{R}$ ($u \neq v$). Once more, because of the singularity of the covariance function along the diagonal, the process can only be understood properly as a random distribution and we can think of the linear statistics (1.5) as a way to regularize it. Note that the Gaussian field G has following spectral representation:

$$G(u) = 2\text{Re} \left\{ \int_0^\infty \frac{\sin(\pi \ell \kappa)}{\sqrt{\kappa}} e^{-2\pi i \kappa u} dB_{\mathbb{C}}(\kappa) \right\}, \quad (1.11)$$

where $B_{\mathbb{C}}$ is a standard complex Brownian motion and may be viewed as a continuous analogue of (1.1).

Our principal interest in this paper will be the regime where $\epsilon_N \rightarrow 0$ depending on N in a suitable way as $N \rightarrow \infty$. In this case, the variance of (1.5) is logarithmically divergent as $N \rightarrow \infty$ so that there is no CLT. However, we shall see in the next section that a Gaussian approximation still holds in a strong sense under suitable conditions on the rate at which $\epsilon_N \rightarrow 0$ and that this suffices to construct the multiplicative chaos measure, as well as to prove the counterparts of conjecture 1.1 for all $q \in \mathbb{N}$.

1.2 Results for the CUE and discussion

To state our results, we recall the setup described in the previous section. We construct the regularized statistic (1.5) by taking the convolution of ϕ with the indicator of the interval $[u - \ell/2, u + \ell/2]$. For the CUE, we will assume that the mollifier ϕ is a smooth function with compact support which is also a probability density. We also assume that the smoothing takes place on scales

$$N^{\alpha-1}\epsilon_N^{-1} = O(N^{-\delta}), \quad N \rightarrow \infty \quad (1.12)$$

for some $\delta > 0$. Note that this is the random matrix theoretic equivalent of the corresponding slow decay condition first formulated in the context of Riemann zeros in [54]. Consider the approximating GMC measure corresponding to (1.5), defined as

$$\mu_{N,\epsilon_N}^\gamma(du) = \exp\left(\tilde{X}_{N,\epsilon}^\gamma(u)\right) du, \quad (1.13)$$

where the centered and renormalized field is

$$\tilde{X}_{N,\epsilon}^\gamma(u) = \gamma(X_{N,\epsilon}(u) - \mathbb{E}(X_{N,\epsilon}(u))) - \frac{\gamma^2}{2}\text{Var}(X_{N,\epsilon}(u)). \quad (1.14)$$

The corresponding limiting measure is the GMC for the field $G(u)$ in (1.11), defined formally by

$$\mu^\gamma(du) = e^{\gamma G(u) - \frac{\gamma^2}{2}\mathbb{E}(G(u)^2)} du. \quad (1.15)$$

Our main result establishes the convergence in the whole sub-critical phase $\gamma < \sqrt{2}$.

Theorem 1.2. *Consider the regularized statistic (1.5) in the mesoscopic regime $0 < \alpha < 1$ and suppose that condition (1.12) holds. Then, for any $\gamma < \sqrt{2}$ and $w \in L^1(\mathbb{R})$ with compact support, the random variable $\mu_{N,\epsilon_N}^\gamma(w)$ converges in distribution as $N \rightarrow \infty$ to $\mu^\gamma(w)$ which is the GMC associated with the log-correlated process (1.11).*

The proof of theorem 1.2 is based on having a sufficiently strong Gaussian approximation of the random field (1.5) combined with the elementary approach to GMC developed by Berestycki in [4]. In fact, our method goes beyond the CUE and would apply to any random field $X_{N,\epsilon}$ (where $\epsilon \geq \delta_N$ may be function of the parameter N) such that for all $u_1, u_2 \in \mathbb{R}$,

$$\text{Cov}(X_{N,\epsilon_1}(u_1), X_{N,\epsilon_2}(u_2)) \sim T_{\epsilon_1,\epsilon_2}(u_1, u_2) \quad \text{as } N \rightarrow \infty,$$

and

$$T_{\epsilon_1,\epsilon_2}(u_1, u_2) \sim \log\left(\min\{|u_1 - u_2|^{-1}, \epsilon_1^{-1}, \epsilon_2^{-1}\}\right) \quad \text{as } |u_1 - u_2| \rightarrow 0,$$

with sufficiently nice error terms. Then, by strong Gaussian approximation, we mean that for any $n \in \mathbb{N}$ and $\epsilon_1, \dots, \epsilon_n \geq \delta_N$ where $\delta_N \rightarrow 0$ as $N \rightarrow \infty$, the Laplace transform satisfies the estimate

$$\mathbb{E}\left[\exp\left(\sum_{k=1}^n t_k \overline{X}_{N,\epsilon_k}(u_k)\right)\right] \sim \exp\left(\frac{1}{2} \sum_{k,j=1}^n t_k t_j T_{\epsilon_j,\epsilon_k}(u_j, u_k)\right). \quad (1.16)$$

uniformly in u_1, \dots, u_n varying in compact subsets of \mathbb{R} , where $\overline{X}_{N,\epsilon}(u) = X_{N,\epsilon}(u) - \mathbb{E}(X_{N,\epsilon}(u))$. In other words, if $(\delta, u) \rightarrow G_\delta(u)$ is the mean-zero Gaussian process on $\mathbb{R}_+ \times \mathbb{R}$ with covariance kernel

$$\mathbb{E}[G_{\delta_1}(u_1)G_{\delta_2}(u_2)] = T_{\delta_1,\delta_2}(u_1, u_2),$$

then the joint exponential moments of the field $\overline{X}_{N,\epsilon}(u)$ are well-approximated by those of $G_\epsilon(u)$. Indeed we formulate our approach in an abstract setting that we believe will have useful applications beyond random matrix theory, where other ‘approximately’ log-correlated Gaussian fields (precisely in the sense of (1.16)) may appear.

Note that the assumption that w is compactly supported in theorem 1.2 is important, as to stay in the mesoscopic regime the values taken by u cannot extend to an interval of length greater than N^α without (1.5) becoming a global statistic. Although ℓ is fixed in theorem 1.2, we now consider the case where $\ell = L(N) \rightarrow \infty$ as $N \rightarrow \infty$. Similarly, the rate at which $L(N) \rightarrow \infty$ is naturally restricted by the mesoscopic scale,

$$L(N)/(N^\alpha \log(N)) \rightarrow 0, \quad N \rightarrow \infty. \quad (1.17)$$

Under this restriction we prove that the moments of the total mass are given by Selberg integrals in the limit $N \rightarrow \infty$.

Theorem 1.3. (*Moments of the total mass and Selberg integrals*) Let $0 < \alpha < 1$ and suppose that the conditions (1.12) and (1.17) are operative. Then for any $r > 0$ and $q \in \mathbb{N}$ such that $\gamma^2 q < 2$, we have

$$\begin{aligned} \lim_{N \rightarrow \infty} L(N)^{-\gamma^2 \binom{q}{2}} \mathbb{E} \left[(\mu_{N, \epsilon_N}^\gamma[0, r])^q \right] &= \lim_{\ell \rightarrow \infty} \lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \ell^{-\gamma^2 \binom{q}{2}} \mathbb{E} \left[(\mu_{N, \epsilon}^\gamma[0, r])^q \right] \\ &= \int_{[0, r]^q} |\Delta(\vec{u})|^{-\gamma^2} du_1 \dots du_q = r^{q + \gamma^2 \binom{q}{2}} \prod_{j=0}^{q-1} \frac{\Gamma(1 - (j+1)\gamma^2/2) \Gamma^2(1 - j\gamma^2/2)}{\Gamma(1 - \gamma^2/2) \Gamma(2 - (q+j-1)\gamma^2/2)}, \end{aligned} \quad (1.18)$$

where $\Delta(\vec{u}) = \prod_{1 \leq j < k \leq q} (u_k - u_j)$.

Theorem 1.3 may be interpreted as a smoothed analogue of conjecture 1.1 and should also be compared to the conjectures given in [54]. The double limit $N \rightarrow \infty$ followed by $\epsilon \rightarrow 0$ in the second equality of (1.18) is referred to generically as the *weak conjecture*, while the more fundamental single limit in the first equality is the *strong conjecture*. Theorem 1.3 demonstrates the equivalence of the strong and weak limits for positive integer $q \in \mathbb{N}$. The full statement of these conjectures actually consists of assertions valid for any $q \in \mathbb{C}$; such generality is beyond the scope of this paper and we restrict ourselves to the integral case of $q \in \mathbb{N}$ throughout. However, it would seem natural to conjecture that Theorem 1.3 holds also for general $q \in \mathbb{C}$ with the appropriate analytic continuation of the Selberg integral in (1.18) given in terms of generalized Barnes functions, *c.f.* equation (49) in [54]. The factor $r^{q + \gamma^2 \binom{q}{2}}$ in (1.18) and the quadratic dependence on q known as *structure exponent* are a consequence of multi-scaling and usually emerge due to the multi-fractal nature of the underlying random measures. It is natural to ask what happens if we take a different interval in (1.4); indeed in [54] the examples $[0, u]$ and $[-L(N), u]$ are studied in detail and the analogous results conjectured there for $q \in \mathbb{N}$ follow easily with our approach. In fact the second example $[-L(N), u]$ behaves similarly to our interval $[u - L(N)/2, u + L(N)/2]$ with $L(N) \rightarrow \infty$. However, it turns out that the random measures corresponding to these examples are *not normalizable* in the usual GMC sense of (1.14). We leave it as an interesting open problem to find a mathematically rigorous way of normalizing the GMC measures corresponding to these interval statistics.

To prove Theorem 1.3 it is clear that the required asymptotics are precisely those in (1.16) with a sufficiently strong uniform control in the variables u_1, \dots, u_q . This uniformity is the essential reason that we perform the smoothing (1.5).

Remark 1.4. If we instead work with the non-smooth quantity (1.4), pointwise asymptotics are a consequence of the known proofs of the Fisher-Hartwig conjecture. For example, Theorem 2.2 in [41] implies that for $u_1 < u_2 < \dots < u_q$ fixed, we have

$$\begin{aligned} \mathbb{E} \left(\exp \left(\sum_{j=1}^q t_j (X_N(u_j) - \mathbb{E}(X_N(u_j))) \right) \right) &= \exp \left(\frac{1}{2} \sum_{j=1}^q t_j^2 \log(N) - \sum_{j < k} t_j t_k \log |e^{iu_k} - e^{iu_j}| \right. \\ &\quad \left. + \sum_{j < k} \frac{t_j t_k}{2} \log \left(|e^{i(u_k - l)} - e^{iu_j}| |e^{i(u_k + l)} - e^{iu_j}| \right) \right) \prod_{j=1}^q |G(1 + \frac{t_j}{2i})|^4 (1 + o(1)) \end{aligned} \quad (1.19)$$

The main difficulty with the asymptotics (1.19) is that the error term is not uniform in varying u_1, \dots, u_q . In fact if $|u_k - u_j| \sim N^{-1}$ a different asymptotic regime is known to emerge involving a Painlevé transcendent, see [13].

This uniformity problem is closely related to the necessity of condition (1.12): without it we enter the transition regime of Fisher-Hartwig singularities, the main difficulty encountered in trying to prove conjecture 1.1 in this way. Our contribution is to show that by smoothing out the field on these scales, we remain in the strong Szegő regime and so the aforementioned uniformity problem is much more tractable. Indeed, for the smoothed field $X_{N, \epsilon_N}(u)$, the Laplace transform can be computed exactly in terms of the *Borodin-Okounkov formula* [8] which allows us a very precise asymptotic analysis. See the book of Simon [64] for a very comprehensive account of this formula.

Finally, let us comment on the universality of our results. It is likely that theorems 1.2 and 1.3 continue to hold for quite general classes of Hermitian random matrices, the main criteria being the validity of the Gaussian asymptotics (1.16). From the point of view of Gaussian multiplicative chaos this is quite natural and one may view the different ensembles of random matrices as alternative ways of regularizing a log-correlated Gaussian field. A sign that we are dealing with a rather fundamental limiting object is its universality and complete independence of the method of regularization. To add some weight to these assertions, we will also prove our results for the *sine process*, the random point process which describes the *local* (microscopic) limit of a wide class of unitary invariant random matrices, [18]. This includes the CUE, but also many of the ensembles for which the CLT (1.6) was recently established [43]. Indeed the CLT (1.6) holds for the sine process itself, as shown by Soshnikov [66].

1.3 Overview of the paper

The paper is organized as follows. In section 1.4, we begin with a brief review on determinantal point processes associated with integrable operators, of which the CUE and sine process are special cases. In section 1.5, we provide a review of some of the recent developments in the theory of Gaussian multiplicative chaos relevant for our needs. Then, we conclude this introduction by constructing a GMC measure associated with regularized counting statistics of the sine process in much the same way we did for the CUE. In section 2.1 and 2.2, we develop the multiplicative chaos theory for random fields with strong Gaussian approximation. In section 2.3, we analyze the covariance structure of the regularization of the log-correlated field $G(u)$ defined in (1.11), proving some preparatory results to apply the general theory to the CUE and sine process. In sections 3.1 and 3.2, we establish the required asymptotics (1.16) for the sine and CUE point processes, using a Riemann-Hilbert problem and the Borodin-Okounkov formula, respectively.

In what follows, $C > 0$ denotes a numerical constant which may change from line to line and we use the notation $a \ll b$ to specify that the quantity $a \leq Cb$. We also define for all $x \in \mathbb{R}$,

$$\log^+(x) = \log(1 \vee |x|).$$

1.4 Determinantal point processes and integrable operators

The aim of this section is to provide, in a general context, a short introduction to the theory of determinantal point processes which focuses on the connection between linear statistics and Fredholm determinants. We also briefly review the concept of integrable operators introduced in [35] and how this relates the Laplace transform of a linear statistic to a Riemann-Hilbert problem.

Let Σ be a Polish space equipped with a Radon measure η . A point configuration $\Upsilon \subset \Sigma$ is a discrete set which is locally finite (i.e. the set $\Upsilon \cap B$ is finite for any compact set $B \subset \Sigma$). A point process is a probability measure on the space of point configurations. This definition can be

made mathematically precise, see for instance [67, 37, 7], and a point process can be described by its intensity measures or correlation functions $\{\rho_n\}_{n=1}^\infty$ which are defined by the formulae:

$$\mathbb{E} \left[\sum_{(\lambda_1, \dots, \lambda_n) \subset \Upsilon} \prod_{k=1}^n f_k(\lambda_k) \right] = \int_{\Sigma^n} \prod_{k=1}^n f_k(x_k) \rho_n(dx_1, \dots, dx_n), \quad (1.20)$$

for any function $f_1, \dots, f_n \in L^\infty(\Sigma \rightarrow \mathbb{R}_+)$ with compact support. Note that the LHS of formula (1.20) consists of a sum over all ordered subsets of the random configuration Υ of size $n \in \mathbb{N}$. A point process is called determinantal if all its intensity measures are of the form

$$\rho_n(dx_1, \dots, dx_n) = \det_{n \times n} [K(x_i, x_j)] \eta(dx_1) \cdots \eta(dx_n). \quad (1.21)$$

The function $K : \Sigma \times \Sigma \rightarrow \mathbb{C}$ is called the correlation kernel. It is obviously not unique, but it encodes the law of the random configuration Υ . There are many interesting examples of determinantal processes coming from probability theory, combinatorics, and mathematical physics such as the eigenvalues of unitary invariant random matrices, free fermions, zeros of Gaussian analytic functions, non-intersecting random walks, uniform spanning trees, random tilings, etc. We refer to the surveys [37, 32, 7] for further examples. Let us just mention the following criterion which goes back to the beginning of the theory, [46], and describes a natural class of correlation kernels.

Theorem 1.5 (Macchi [46], Soshnikov [67]). *If a kernel K determines a self-adjoint integral operator acting on $L^2(\Sigma)$ which is locally trace-class, then K defines a determinantal point process if and only if its spectrum is contained in $[0, 1]$.*

In the following, we shall assume that the kernel K is a continuous function on $\Sigma \times \Sigma$ and satisfies the hypothesis of theorem 1.5. In this case, this kernel defines an operator, also denoted by K , which is locally trace-class if and only if for any compact set $B \subseteq \Sigma$,

$$\text{Tr } K = \int_B K(x, x) dx < \infty;$$

see [65, Theorem 2.12]. We let Υ be the point configuration of the determinantal process with kernel K and for any function $\varphi \in L^\infty(\Sigma \rightarrow \mathbb{R}_+)$, we denote

$$K_\varphi(x, x') = \sqrt{\varphi(x)} K(x, x') \sqrt{\varphi(x')}. \quad (1.22)$$

The condition that the function $\varphi \geq 0$ is not necessary but rather convenient. In particular, this implies that the operator K_φ is also self-adjoint, non-negative, and it is trace class if

$$\text{Tr } K_\varphi = \int_\Sigma K(x, x) \varphi(x) dx = \mathbb{E} \left[\sum_{\lambda \in \Upsilon} \varphi(\lambda) \right] < \infty. \quad (1.23)$$

Note that this condition holds if for instance φ has compact support. The last equality in (1.23) follows from the definition of the first intensity measure and the function $x \mapsto K(x, x)$ is called the density of the point process Υ . The reason to consider the kernel (1.22) is that, using formulae (1.20) and (1.21), it is a simple combinatorial exercise to show that if K_φ is trace-class, then for any $t \geq 0$,

$$\mathbb{E} \left[\prod_{\lambda \in \Upsilon} (1 + t\varphi(\lambda)) \right] = \det[\text{I} + tK_\varphi]_{L^2(\Sigma)}, \quad (1.24)$$

where the RHS is a Fredholm determinant; c.f. [37]. In particular, taking the usual logarithm, we obtain

$$\log \mathbb{E} \left[\prod_{\lambda \in \Upsilon} (1 + t\varphi(\lambda)) \right] = \text{Tr } \log(\text{I} + tK_\varphi), \quad (1.25)$$

and this function is differentiable for all $t > 0$:

$$\frac{d}{dt} \text{Tr} \log(\mathbf{I} + tK_\varphi) = \text{Tr} \left[\frac{K_\varphi}{\mathbf{I} + tK_\varphi} \right].$$

Hence, if we define $L_t := \frac{K_\varphi}{\mathbf{I} + tK_\varphi}$, this implies that

$$\log \mathbb{E} \left[\prod_{\lambda \in \Upsilon} (1 + \varphi(\lambda)) \right] = \int_0^1 \text{Tr}[L_t] dt. \quad (1.26)$$

For instance, taking $\varphi(x) = \mathbf{1}_{x \in B}$ for some compact subset $B \subseteq \Sigma$, one can investigate the distribution of the random variable $|\Upsilon \cap B|$ and in particular the probability that there are no points in the set B . More generally, if $h \in L^\infty(\Sigma \rightarrow \mathbb{R}_+)$, taking $\varphi(x) = e^{h(x)} - 1$, this gives an explicit formula for the moment generating function or Laplace transform of the linear statistic $\sum_{\lambda \in \Lambda} h(\lambda)$. As the density of the point process converges to infinity, this reduces the question about the statistical properties of the random variable $\sum_{\lambda \in \Upsilon} h(\lambda)$ to a question about the asymptotics of the resolvent operator L_t . There is a special class of determinantal processes, those for which the correlation kernel gives rise to an integrable operator, which are particularly interesting because computing the resolvent L_t turns out to be equivalent to solving a Riemann-Hilbert problem; see proposition 1.6 below. In particular, it allows to use the so-called *non-linear steepest descent* method introduced in [19] to obtain the asymptotics of formula (1.26). The theory of integrable operators and the auxiliary Riemann-Hilbert problem originates in the context of statistical field theory, [35], but this approach has also been used to answer different types of questions about the statistics of eigenvalues of unitary invariant matrix ensembles. For instance, one can find a proof of the Strong Szegő limit theorem in [16] and, in [6], the author extended Deift's method to investigate a transition for smooth linear statistics of the so-called thinned CUE and thinned sine process. To our best knowledge, this was the first instance where this technique was applied to mesoscopic linear statistics. In this paper, we will use an analogous method to derive the necessary estimates to construct a multiplicative chaos measure which arises naturally from the sine process. In particular, we will make use of the following result from [35], see also [16].

Theorem 1.6. *Suppose that Σ is a closed (oriented) curve on the Riemann sphere. Let Υ be a determinantal process on Σ with Hermitian correlation kernel of the form*

$$K(z, z') = -\frac{f(z)^* g(z')}{2\pi i(z - z')}, \quad (1.27)$$

where $f : \Sigma \rightarrow \mathbb{C}^k$ and $g : \Sigma \rightarrow \mathbb{C}^k$ are continuously differentiable functions so that $f(z)^* g(z) = 0$ for all $z \in \Sigma$. If $\varphi : L^\infty(\Sigma \rightarrow \mathbb{R})$ is a test function so that both $\sqrt{\varphi}f, \sqrt{\varphi}g \in L^2(\Sigma)$, then we have

$$\log \mathbb{E} \left[\prod_{\lambda \in \Upsilon} (1 + \varphi(\lambda)) \right] = \frac{-1}{2\pi i} \int_0^1 \int_{\mathbb{R}} \left(\frac{d\sqrt{\varphi}F_t}{dx}(x) \right)^* \left(\sqrt{\varphi(x)}G_t(x) \right) dx dt \quad (1.28)$$

where $F_t = m_+ f$ and $G_t = (m_+^{-1})^* g$ and the matrix m is the (unique) solution of the Riemann-Hilbert problem:

- $m(z)$ is analytic on $\mathbb{C} \setminus \Sigma$.
- If we let $v = \mathbf{I} + t\varphi f g^*$, then $m(z)$ satisfies the jump condition:

$$m_+(z) = m_-(z)v(z), \quad z \in \Sigma \quad (1.29)$$

- $m(z) \rightarrow \mathbf{I}$ as $z \rightarrow \infty$ in $\mathbb{C} \setminus \Sigma$.

In practice, $\Sigma = \{|z| = 1\}$ for the CUE, or $\Sigma = \mathbb{R}$ for the sine process and the eigenvalue processes coming from unitary invariant ensembles. Moreover, the correlation kernels of these processes all give rise to integrable operators. According to formula (1.27), we may choose for the CUE:

$$f_{\mathcal{U}_N}(z) = \begin{pmatrix} z^{N+1} \\ 1 \end{pmatrix} \quad \text{and} \quad g_{\mathcal{U}_N}(z) = \begin{pmatrix} z^{N+1} \\ -1 \end{pmatrix}$$

and for the sine process:

$$f(x) = \begin{pmatrix} e^{i\pi Nx} \\ e^{-i\pi Nx} \end{pmatrix} \quad \text{and} \quad g(x) = \begin{pmatrix} e^{i\pi Nx} \\ -e^{-i\pi Nx} \end{pmatrix}.$$

For this example, the Riemann-Hilbert problem (1.29) is solved in section 3.1 for a large class of analytic test function φ ; see in particular proposition 3.1 for the asymptotics of the solution. As a last comment about universality. If one consider the eigenvalues of an $N \times N$ Hermitian random matrix sampled according to the weight $e^{-NV(H)}$ for a real-analytic external field $V : \mathbb{R} \rightarrow \mathbb{R}$, then, by the Christoffel-Darboux formula, the correlation kernel of the eigenvalue process is also integrable with

$$f_V(x) = \begin{pmatrix} \pi_N(x) \\ -2\pi i \gamma_{N-1}^2 \pi_{N-1}(x) \end{pmatrix} e^{-NV(x)/2} \quad \text{and} \quad g_V(y) = \begin{pmatrix} -2\pi i \gamma_{N-1}^2 \pi_{N-1}(x) \\ \pi_N(x) \end{pmatrix} e^{-NV(x)/2}. \quad (1.30)$$

where π_N and π_{N-1} are the monic polynomials with respect to the measure $e^{-NV(x)}dx$ on \mathbb{R} of degree N and $N-1$ respectively and

$$\gamma_{N-1}^{-2} = \int \pi_{N-1}(x)^2 e^{-NV(x)} dx.$$

Observe that, f_V is exactly the first row of the solution Y_N of the orthogonal polynomial Riemann-Hilbert problem whose solution is derived in detail in [18]. In particular from their results, one can extract the universal oscillatory behavior of the function f_V and g_V in the bulk which indicates strongly that the approach we present for the sine process could be generalized and would provide a way to show that the limiting chaos measure has the same law for a large class of potentials V . However, turning this heuristic into a rigorous computation is rather technical and we leave it as an open problem for future work.

1.5 Gaussian Multiplicative Chaos

In this section, we briefly review the theory of Gaussian Multiplicative Chaos (GMC) with respect to the Lebesgue measure on a compact subset $\mathfrak{A} \subset \mathbb{R}^d$. This theory which originates in the work of Mandelbrot and Kahane [47, 48, 39, 38] aims at defining the exponential of a log-correlated random field G , denoted formally by

$$\nu^\gamma(dx) = e^{\gamma G(x)} dx. \quad (1.31)$$

The original motivation to study such an object goes back to the work of Kolmogorov who proposed that the measure ν^γ should describe the energy dissipation in a turbulent fluid; c.f. [56] for a modern reference. Another motivation comes from the fact that ν^γ can be interpreted as the Boltzmann-Gibbs measure associated with the *random Hamiltonian* G . Then, this measure describes the equilibrium configuration of a particle in a very rough landscape and $\gamma > 0$ plays the role of the inverse-temperature and is usually called the intermittency parameter. In fact, sampling from the measure ν^γ gives information about the points where the field G takes unusually high values known as γ -thick points, see [33]. In particular, there exists a critical value of γ above which no such points exist and the measure (1.31) needs to be renormalized in a different way and becomes purely atomic. This is known as the freezing transition in the theory of spin glasses and it has been observed that the behavior of ν^γ at criticality is related to the law of the maximum of

the field G [23, 25]. Recently, there have also been intense developments in the case where G is the Gaussian Free Field associated to a domain in the complex plane. Then, the chaos measure (1.31) is known as the Liouville measure and it has been one of the key inputs in a program aiming at giving a mathematically rigorous construction of Liouville quantum gravity and Liouville quantum field theory, c.f. the recent results of [15] and [49]. The latter paper is concerned with developing an important program of *imaginary geometries* and *Liouville quantum gravity and the Brownian map* which aim at proving that the Liouville measures are central objects which arise, for instance to describe the scaling limit of random planar maps. In addition various KPZ relations have been established [3, 20, 57]. For a more detailed introduction and further references to these topics, we refer to the lecture notes [30, 5, 59] or the survey [58].

The dots : $G(x)$: in (1.31) refer to the fact that the field G is merely a random distribution, so that its exponential is not defined in the usual way. Usually, the random measure ν^γ is defined by first regularizing the field G in some way and then by taking a limit as the regularization tends to zero. There have been many important developments in understanding this procedure and the so-called subcritical case is now well understood for Gaussian regularisations [38, 60, 63, 4]. To be specific, let G be a Gaussian process on \mathfrak{A} with a covariance kernel:

$$T(x, y) = -\log |x - y| + g(x, y)$$

where the function $g : \mathfrak{A}^2 \rightarrow \mathbb{R} \cup \{-\infty\}$ is continuous as an extended function and there exists a constant $C > 0$ so that for all $x, y \in \mathfrak{A}$,

$$g(x, y) - \log^+ |x - y| \leq C. \quad (1.32)$$

We introduce this general setting since our main example is a stationary Gaussian process on \mathbb{R} with covariance kernel Q given by (1.10). Note that one usually assumes that g is a continuous and bounded function, but these assumptions can be relaxed without changing the general theory, as long as a condition such as (1.32) holds. Because of the singularity of the kernel T on the diagonal, G is not defined pointwise and needs to be interpreted as a random generalized function. Formally $G = \{G(f)\}_{f \in C(\mathfrak{A})}$ is a Gaussian process with covariance structure:

$$\mathbb{E}[G(f)G(g)] = \iint_{\mathbb{R}^2} f(x)g(y)T(x, y)dx dy.$$

In general, the definition of G can be extended to a more general class of test functions than $C(\mathfrak{A})$, or even to certain classes of measures. To define the exponential measure (1.31), one can consider a regularization of the process G coming from the convolution with an approximate delta function. This approach was introduced by Robert and Vargas in [60] and developed further by Berestycki in [4]. Namely, given $\epsilon > 0$ and a mollifier ϕ (i.e. a sufficiently smooth and light-tailed probability density function on \mathbb{R}), we define

$$G_{\phi, \epsilon} = G \star \phi_\epsilon \quad \text{where} \quad \phi_\epsilon(x) = \epsilon^{-1} \phi(x/\epsilon),$$

and for any $\gamma > 0$,

$$\nu_{T, \epsilon}^\gamma(dx) = \exp \left(\gamma G_\epsilon(x) - \frac{\gamma^2}{2} \mathbb{E}[G_\epsilon(x)^2] \right) dx. \quad (1.33)$$

Then, for any function $w \in L^1(\mathfrak{A} \rightarrow \mathbb{R}_+)$, if $\gamma^2 < 2d$, we let

$$\nu_T^\gamma(w) = \lim_{\epsilon \rightarrow 0} \nu_{T, \epsilon}^\gamma(w). \quad (1.34)$$

The random measures ν_T^γ are called the multiplicative chaos measures associated to the Gaussian process with covariance kernel T . The major achievements of the GMC theory are that the limit

(1.34) exists in probability (and almost surely in certain cases) and that it does not depend on ϕ for a large class of mollifiers. Moreover, it is non-trivial if and only if $\gamma^2 < 2d$. In the critical ($\gamma = \sqrt{2d}$) and supercritical ($\gamma^2 > 2d$) regimes, one needs different normalizations than (1.33) to make sense of ν_T^γ in a non-trivial way, c.f. [58, Section 6] and reference therein. In this paper, we will focus on the subcritical regime, in which case by [4], the limit (1.34) holds in $L^1(\mathbb{P})$ and the normalization is such that

$$\mathbb{E}[\nu_T^\gamma(w)] = \int w(u) du.$$

Moreover, for any $q \in \mathbb{N}$ such that $q\gamma^2 < 2$, it is not difficult to show that

$$\mathbb{E}[\nu_T^\gamma(w)^q] := \lim_{\epsilon \rightarrow 0} \mathbb{E}[\nu_{T,\epsilon}^\gamma(w)^q] = \int_{\mathfrak{A}^q} \exp\left(\gamma^2 \sum_{1 \leq j < k \leq q} T(u_j, u_k)\right) \prod_{k=1}^q w(u_k) du_k. \quad (1.35)$$

In particular, in dimension $d = 1$, by a change of variables, this implies that for any $x_0 \in \mathbb{R}$,

$$\mathbb{E}[\mu_T^\gamma([x_0 - \frac{r}{2}, x_0 + \frac{r}{2}])^q] = r^{\xi(q)} \int_{[0,1]^q} \prod_{1 \leq j < k \leq q} |u_j - u_k|^{-\gamma^2} e^{\gamma^2 g(x_0 + r(u_j - \frac{1}{2}), x_0 + r(u_k - \frac{1}{2}))} \prod_{k=1}^q du_k,$$

where $\xi(q) = q - \gamma^2 \frac{q(q-1)}{2}$, so that as $r \rightarrow \infty$, we have

$$\mathbb{E}[\mu_T^\gamma([x_0 - \frac{r}{2}, x_0 + \frac{r}{2}])^q] \sim r^{\xi(q)} \mathcal{S}(q; \gamma^2/2) e^{\gamma^2 (\frac{q}{2}) g(x_0, x_0)} \quad \text{as } r \rightarrow 0,$$

where

$$\mathcal{S}(n; \tilde{\gamma}) := \int_{[0,1]^n} \prod_{i \neq j} |u_i - u_j|^{-\tilde{\gamma}} \prod_{k=1}^n du_k = \prod_{j=0}^{n-1} \frac{\Gamma(1 + j\tilde{\gamma})^2 \Gamma(1 + \tilde{\gamma} + j\tilde{\gamma})}{\Gamma(2 + (n + j - 1)\tilde{\gamma}) \Gamma(1 + \tilde{\gamma})} \quad (1.36)$$

is a Selberg integral. The quadratic function $\xi(q)$ is called the **structure exponent** and it describes the multi-fractal properties of the random measure μ_T^γ , c.f. [58, section 2.3]. Finally, since the Selberg integral converges if and only if $\Re\{n\tilde{\gamma}\} < 1$, this shows that the condition $q\gamma^2 < 2$ in (1.35) is sharp.

Let us conclude this section by stating a result about the convergence of the GMC measure in the so-called L^2 -phase ($\gamma^2 < d$). In particular, we will need this result in section 2.1 in order to identify the law of the multiplicative chaos measure coming from counting statistics of the CUE or sine process. In addition, the strategy of the proof will be re-used and generalized in sections 2.1 and 2.2 to construct non-Gaussian multiplicative chaos measures.

Proposition 1.7. *Let q be an even integer and $\gamma > 0$ so that $q\gamma^2 < 2d$. For any $w \in L^1(\mathfrak{A})$, the random variable $\nu_{Q,\epsilon}^\gamma(w)$ given by (1.33) converges as $\epsilon \rightarrow 0$ in $L^q(\mathbb{P})$ to a random variable $\nu_Q^\gamma(w)$ which does not depend on the mollifier ϕ subject to the conditions that ϕ is smooth and $\phi \in \mathcal{D}_\alpha$ for some sufficiently large $\alpha > 0$, (1.43).*

Proof. This follows from a standard argument; c.f. for instance the proof of [59, Theorem 2.3]. First, given a mollifier ϕ , one can prove that $\nu_{T,\epsilon}^\gamma(w)$ is a Cauchy sequence in $L^q(\mathbb{P})$. When q is an even integer, this just boils down to checking that

$$\mathbb{E}[G_{\phi,\epsilon}(x)G_{\phi,\epsilon'}(x')] \rightarrow T(x, x') \quad (1.37)$$

as $\epsilon, \epsilon' \rightarrow 0$ in such a way that we can apply the dominated convergence theorem. To prove that the limit does not depend on ϕ , if $\nu_{T,\epsilon}^\gamma$ is the measure associated to the Gaussian process $G_{\psi,\epsilon}$ for a second mollifier ψ , then it suffices to show that

$$\lim_{\epsilon \rightarrow 0} \mathbb{E}[|\nu_{T,\epsilon}^\gamma(w) - \nu_{T,\epsilon}^\gamma(w)|^q] = 0.$$

This limit follows in a similar fashion by checking that as $\epsilon \rightarrow 0$,

$$\mathbb{E}[G_{\phi,\epsilon}(x)G_{\phi,\epsilon}(x')] \rightarrow T(x, x'). \quad \square$$

Remark 1.8. *In what follows, the condition (1.37) is replaced by the assumption 2.3. In fact, by going carefully through the proof in [4], it is not difficult to check that if*

$$T_{\epsilon,\delta}(x, x') = \mathbb{E}[G_{\phi,\epsilon}(x)G_{\phi,\delta}(x')]$$

satisfies the assumption 2.3, then for any $\gamma < \sqrt{2d}$, $\nu_{Q,\epsilon}^\gamma(w)$ converges in $L^1(\mathbb{P})$ to $\nu_Q^\gamma(w)$ as $\epsilon \rightarrow 0$. Moreover, for the stationary Gaussian process on \mathbb{R} with covariance kernel Q given by (1.10) which arises in theorem 1.2 for the mesoscopic CUE, as well as in theorem 1.9 for the sine process, it is proved in section 2.3 that the assumption 2.3 holds for any mollifier $\phi \in \mathcal{D}_\alpha$ for any $\alpha > 0$.

1.6 Results for the sine process

In this section, we show how the results of section 2 apply to regularized counting statistics of the sine process. The results are analogous to those stated for the CUE in the introduction. Hence, we expect that similar results hold for well-behaved unitary invariant ensembles as well, but we leave the task of proving the sufficient asymptotics open for a future project. The sine process Λ_N is the determinantal point process on \mathbb{R} with correlation kernel

$$K_N(u, v) = \frac{\sin(\pi N(u - v))}{\pi(u - v)}. \quad (1.38)$$

This is a translation invariant point process whose density is N times the Lebesgue measure on \mathbb{R} . Recall that, given a mollifier ϕ , we denote $\phi_\epsilon(x) = \epsilon^{-1}\phi(x/\epsilon)$ and $\chi_u(x) = \pi \mathbb{1}_{|x-u| \leq \ell/2}$ for all $u, x \in \mathbb{R}$. Let us consider the linear statistics:

$$X_{N,\epsilon}(u) = \sum_{\lambda \in \Lambda_1} (\chi_u * \phi_\epsilon)(\lambda/N). \quad (1.39)$$

As $N \rightarrow \infty$, the random variable (1.39) describes the number of eigenvalues in a mesoscopic box in the bulk of the GUE, or say of another unitary invariant ensemble, and to see the parallel with the CUE, note that the scaling property of the sine kernel implies that

$$X_{N,\epsilon}(u) \stackrel{d}{=} \sum_{\lambda \in \Lambda_N} \chi_u * \phi_\epsilon(\lambda). \quad (1.40)$$

We will focus on the strong regime where $\epsilon = \epsilon_N \rightarrow 0$ as $N \rightarrow \infty$ and, viewing $X_{N,\epsilon}$ as an asymptotically Gaussian field on \mathbb{R} , we will construct its chaos measure in the subcritical phase. The advantage of working with the random variables (1.39) instead of the RHS of (1.40) is that it introduces a natural coupling which allows us to obtain a stronger mode of convergence than for the CUE. For technical reasons, we consider the following class of real-analytic mollifiers.

Assumption 1.1. *Suppose that the function ϕ is analytic in $|\Im z| < \epsilon$ and so that $|\Im \phi| < \pi/\ell$ in this strip. We also assume that $\phi \geq 0$ on \mathbb{R} , $\int_{\mathbb{R}} \phi(x)dx = 1$, and $\sup \{\|\phi\|_{L^1(\mathbb{R}+is)} : s \leq \epsilon/2\} = C_\phi < \infty$.*

For any $\gamma > 0$, we define

$$\tilde{X}_{N,\epsilon}^\gamma(u) = \gamma X_{N,\epsilon}(u) - \gamma \mathbb{E}[X_{N,\epsilon}(u)] - \frac{\gamma^2}{2} \text{Var}[X_{N,\epsilon}(u)] \quad (1.41)$$

and we consider the random measure

$$\mu_{N,\epsilon}^\gamma(du) = \exp\left(\tilde{X}_{N,\epsilon}^\gamma(u)\right) du. \quad (1.42)$$

For any $\alpha \geq 0$, we let

$$\mathcal{D}_\alpha = \left\{ \phi \in L^1 \cap L^2(\mathbb{R}) : \phi \geq 0, \int_{\mathbb{R}} \phi(x) dx = 1 \text{ and } \int_{\mathbb{R}} |x|^\alpha \phi(x) dx < \infty \right\} \quad (1.43)$$

and $\mathcal{D} = \bigcup_{\alpha > 0} \mathcal{D}_\alpha$.

Finally recall that G is the stationary Gaussian field on \mathbb{R} with zero mean and covariance kernel (1.10), and that ν_Q^γ denotes the GMC measure corresponding to G .

Theorem 1.9. *Let $w \in L^1(\mathbb{R})$, $\phi \in \mathcal{D}$ be a function which satisfies the assumption 1.1, and let ϵ_N be a sequence which converges to 0 as $N \rightarrow \infty$ in such a way that $\epsilon_N \geq N^{-1}(\log N)^\beta$ for some $\beta > 1$. For any $0 < \gamma < \sqrt{2}$, $\mu_{N, \epsilon_N}^\gamma(w)$ converges in $L^1(\mathbb{P})$ as $N \rightarrow \infty$ to a random variable $\mu^\gamma(w)$ which has the same law as $\nu_Q^\gamma(w)$.*

This shows that, in the subcritical phase, the law of the random measure μ^γ does not depend on the mollifier ϕ and it is the same GMC measure as for the CUE. The proof of theorem 1.9 is given at the end of section 2.2 and is a direct consequence of our general result, theorem 2.6. In fact, our asymptotics are sufficiently strong to strengthen the convergence of the measure $\mu_{N, \epsilon}^\gamma$ when the parameter γ is sufficiently small. In particular, motivated by the conjectures of [54], beyond the L^2 -phase, we establish the convergence of all the existing moments of the multiplicative chaos measure μ^γ .

Theorem 1.10. *Under the same assumptions as theorem 1.9, if q is an even integer and $\gamma \geq 0$ so that $q\gamma^2 < 2$, then the random variable $\mu_{N, \epsilon}^\gamma(w)$ converges in $L^q(\mathbb{P})$ to $\mu^\gamma(w)$. Moreover, for any $q \in \mathbb{N}$ and $\gamma \geq 0$ such that $q\gamma^2 < 2$, we have*

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\mu_{N, \epsilon}^\gamma(w)^q \right] = \int_{\mathbb{R}^q} \exp \left(\gamma^2 \sum_{1 \leq j < k \leq q} Q(u_j - u_k) \right) \prod_{k=1}^q w(u_k) du_k. \quad (1.44)$$

Theorem 1.10 is an easy variation of theorem 2.1. The details of the proof (in particular, how to deduce the necessary assumptions from the Riemann-Hilbert asymptotics of section 2.3) are given just after we state our main asymptotic result which is valid for the Laplace transform of linear statistics for a broad class of real-analytic test functions.

Assumption 1.2. *Let $0 < \alpha < \pi$. Suppose that h is a function which is analytic and satisfies*

$$|\Im h(z)| < \alpha,$$

in a strip $|\Im z| < \delta$. We also assume that $h : \mathbb{R} \rightarrow \mathbb{R}$, that $h' \in L^1(\mathbb{R})$, and that the following constants are finite:

$$C_\infty = \sup \left\{ \exp |h(z)| : |\Im z| \leq \delta/2 \right\} \text{ and } C_1 = \sup \left\{ \|h\|_{L^\infty(\mathbb{R}+is)} \vee \|h\|_{L^1(\mathbb{R}+is)} : s \leq \delta/2 \right\}. \quad (1.45)$$

Proposition 1.11. *If h is a test function which satisfies the assumption 1.2, then*

$$\log \mathbb{E} \left[\exp \left(\sum_{\lambda \in \Lambda_N} h(\lambda) \right) \right] = N \int h(x) dx + \frac{1}{2} \|h\|_{H^{1/2}(\mathbb{R})}^2 + O_{N \rightarrow \infty} \left(\frac{C_\infty^6 C_1^2}{\delta^{3/2} |\sin \alpha|} e^{-\pi \delta N} \right). \quad (1.46)$$

Note that the leading terms correspond to the mean and variance of the linear statistic $\sum_{\lambda \in \Lambda_N} h(\lambda)$. In particular, (1.46) implies that such linear statistics satisfy a CLT. This is a classical result, but we need to take extra care to control the error term uniformly, especially because we will consider N -dependent test functions. Specifically, we can consider any regime where $\delta(N) \rightarrow 0$ as $N \rightarrow \infty$ almost as fast as N^{-1} (i.e. up to the critical regime where the Gaussian approximation fails). The proof of proposition 1.11 is given in section 3.1 and it relies on the representation of proposition 1.6

for the LHS of (1.46) and applying the Deift-Zhou steepest descent method to the corresponding Riemann-Hilbert problem.

Proof of theorem 1.9. For any $\delta > 0$, define

$$\triangle_n(\delta) = \{\epsilon \in \mathbb{R}^n : \epsilon_1 > \dots > \epsilon_n > \delta\}. \quad (1.47)$$

Let $\delta_N = N^{-1}(\log N)^\beta$, ϕ be a function which satisfies the assumption 1.1, and define for any $\mathbf{t} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^n$ and $\epsilon \in \triangle_n(\delta_N)$,

$$h_{\mathbf{t}, \mathbf{u}, \epsilon}(z) = \pi \sum_{k=1}^n t_k \int_{\mathbb{R}} \mathbb{1}_{|x-u_k| \leq \ell/2} \phi_{\epsilon_k}(z-x) dx.$$

This function is analytic in the strip $|\Im z| \leq \epsilon \delta_N$ and we claim that it satisfies the assumption 1.2 with

$$C_\infty(h_{\mathbf{u}, \epsilon}) = e^{C_\phi |\mathbf{t}|} \quad \text{and} \quad C_1(h_{\mathbf{u}, \epsilon}) = C_\phi |\mathbf{t}|(\ell \vee 1),$$

where $|\mathbf{t}| = |t_1| + \dots + |t_n|$. To check this assumption, we can use the bounds:

$$\left| \int_{\mathbb{R}} \mathbb{1}_{|x-u| \leq \ell/2} \phi_\epsilon(z-x) dx \right| \leq \int_{\mathbb{R}} |\phi(x + \Im z)| dx$$

and

$$\iint_{\mathbb{R} \times \mathbb{R}} \mathbb{1}_{|x-u| \leq \ell/2} |\phi_\epsilon(t-x+is)| dx dt \leq \ell \int_{\mathbb{R}} |\phi(x+is)| dx,$$

which hold for any $u \in \mathbb{R}$ and $\epsilon > 0$. This implies that we can apply proposition 1.11. Moreover, since the sine process has constant density N on \mathbb{R} , the leading term in formula (1.46) corresponds to the expected value of the linear statistic $\sum_{\lambda \in \Lambda_N} h_{\mathbf{u}, \epsilon}(\lambda)$, and by definition of the $H^{1/2}$ Gaussian noise Ξ (see the appendix A), the second order term corresponds to the variance of the Gaussian random variable $\Xi(h_{\mathbf{u}, \epsilon})$. Thus, we get

$$\log \mathbb{E} \left[\exp \left(\sum_{\lambda \in \Lambda_N} h_{\mathbf{u}, \epsilon}(\lambda) \right) \right] = \mathbb{E} \left[\sum_{\lambda \in \Lambda_N} h_{\mathbf{u}, \epsilon} \right] + \frac{\mathbb{E} [\Xi(h_{\mathbf{u}, \epsilon})^2]}{2} + O_{N \rightarrow \infty} \left(\delta_N^{-3/2} (|\mathbf{t}| \ell)^2 e^{6C_\phi |\mathbf{t}| - \pi \delta_N N} \right).$$

By definition of the random field (1.39) and the scaling property (1.40), we see that the linear statistics

$$\sum_{\lambda \in \Lambda_N} h_{\mathbf{u}, \epsilon}(\lambda) \stackrel{d}{=} \sum_{k=1}^n X_{N, \epsilon_k}(u_k),$$

and using the representation (A.15), we also have $\Xi(h_{\mathbf{u}, \epsilon}) \stackrel{d}{=} \sum_{k=1}^n G_{\phi, \epsilon_k}(u_k)$. Hence, in the regime where $|\mathbf{t}|$ and $\ell > 0$ are independent of the parameter N , this implies that for any $\beta > 1$ and $\epsilon \in \triangle_n(\delta_N)$,

$$\log \mathbb{E} \left[\exp \left(\sum_{k=1}^n t_k X_{N, \epsilon_k}(u_k) \right) \right] = \frac{1}{2} \sum_{k,j=1}^n t_k t_j \mathbb{E} [G_{\phi, \epsilon_j}(u_j) G_{\phi, \epsilon_k}(u_k)] + O_{N \rightarrow \infty} \left(e^{-(\log N)^\beta} \right), \quad (1.48)$$

uniformly for all $\mathbf{u} \in \mathbb{R}^n$. In particular, this immediately shows that the random field $u \mapsto X_{N, \epsilon}(u)$ satisfies the assumptions 2.1 and 2.4 for all $q \in \mathbb{N}$. Moreover, if the mollifier $\phi \in \mathcal{D}$, by corollary 2.14, the assumption 2.3 holds too. Consequently, by theorem 2.1, we obtain the convergence of the moments, formula (1.44). Then, if $w \in L^1(\mathbb{R})$ and q is an even integer, to prove convergence in $L^q(\mathbb{P})$ of the random variable $\mu_{N, \epsilon}^\gamma(w)$, we use proposition 2.5. Namely, if $\gamma^2 q < 2$, by proposition 1.7

and remark 1.8, the random variable $\mu_\epsilon^\gamma(w)$ constructed in proposition 2.5 converges in $L^q(\mathbb{P})$ as $\epsilon \rightarrow 0$ to a random variable $\mu^\gamma(w)$ which has the same law as $\nu_Q^\gamma(w)$. In other words, we have

$$\lim_{\epsilon \rightarrow 0, N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\epsilon}^\gamma(w) - \mu^\gamma(w) \right|^q \right] = 0.$$

On the other hand, by lemma 2.4 and the triangle inequality, this implies that

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\epsilon_N}^\gamma(w) - \mu^\gamma(w) \right|^q \right] = 0$$

which concludes the proof of theorem 1.9. \square

Remark 1.12. *In principle, it could be possible to work with more general mollifier by using an argument analogous to the one in [6]. It consists in constructing an N -dependent approximation $\phi^{(N)}$ of the mollifier ϕ which is real-analytic so that we can solve the Riemann-Hilbert for $\phi^{(N)}$, c.f. lemma 3.1, and argue that the Laplace transform of the two regularization are sufficiently close as $N \rightarrow \infty$.*

Acknowledgements

We thank Christian Webb for his interest in our work, for useful discussions and carefully reading the first draft of this paper which helped fix typos and mistakes in the proofs and greatly improved the presentation. G. L. thanks Juhan Aru for stimulating discussions about Gaussian multiplicative chaos. N. S. would like to thank Yan Fyodorov for suggesting to study the conjectures of [54] in the context of random matrix theory.

2 Proof of the main results

Even though the applications discussed in this paper are concerned with random processes defined on \mathbb{R} , we will formulate our convergence results in an abstract setting under some general assumptions. Let \mathfrak{A} be a compact ball in \mathbb{R}^d , $d \geq 1$. In certain cases, like for the sine process, one might also consider the case where \mathfrak{A} is not compact, this requires only slight modifications of our proof. Like in section 1.5, we consider a real-valued Gaussian process G on \mathfrak{A} with a covariance kernel:

$$T(x, y) = -\log|x - y| + g(x, y),$$

where the function $g : \mathfrak{A}^2 \rightarrow \mathbb{R} \cup \{-\infty\}$ is continuous and such that there exists a constant $C > 0$ so that for all $x, y \in \mathbb{R}$,

$$g(x, y) - \log^+ |x - y| \leq C.$$

We also consider a family of real-valued random fields $X_{N,\epsilon}(u)$ defined on \mathfrak{A} which are centered, depend on two parameters $N, \epsilon > 0$, and behave asymptotically like the Gaussian process G . Specifically, we should assume that for any $N > 0$, $(u, \epsilon \mapsto X_{N,\epsilon}(u))_{u \in \mathfrak{A}, \epsilon > 0}$ are random processes defined on the same probability space and that they satisfy the assumptions 2.1 – 2.4 below. For any $\gamma \in \mathbb{R}$, we consider the normalized process

$$\tilde{X}_{N,\epsilon}^\gamma(u) = \gamma X_{N,\epsilon}(u) - \frac{\gamma^2}{2} \mathbb{E} [X_{N,\epsilon}(u)^2], \quad (2.1)$$

and we aim at constructing the limit of the random measure

$$\mu_{N,\epsilon}^\gamma(du) = \exp \left(\tilde{X}_{N,\epsilon}^\gamma(u) \right) du$$

when $\epsilon_N \rightarrow 0$ as $N \rightarrow \infty$ sufficiently slowly. To begin with, in section 2.1, we shall prove that $\mu_{N,\epsilon}^\gamma$ converges to a GMC measure in the L^2 -phase ($\gamma < \sqrt{d}$) and compute the limit of its moments in view of proving theorem 1.3. Then, in section 2.2, we tackle the more challenging task to show that $\mu_{N,\epsilon}^\gamma$ converges in the whole subcritical regime ($\gamma < \sqrt{d}$), by generalizing the method of Berestycki, [4], in an asymptotically Gaussian setting.

2.1 Convergence of the multiplicative chaos measure in the L^2 -phase

Assumption 2.1 (Finite-dimensional convergence in the weak regime). *For any given $\epsilon > 0$, we have*

$$\lim_{N \rightarrow \infty} \mathbb{E} [X_{N,\epsilon}(u) X_{N,\epsilon}(v)] = T_{\epsilon,\delta}(u, v),$$

and the field $u \mapsto X_{N,\epsilon}(u)$ converges in the sense of finite-dimensional distributions to a mean-zero Gaussian process G_ϵ with covariance structure:

$$T_{\epsilon,\delta}(u, v) = \mathbb{E} [G_\epsilon(u) G_\delta(v)], \quad (2.2)$$

for any $u, v \in \mathfrak{A}$ and $\epsilon, \delta > 0$.

In the context of random matrix theory described in the introduction, the assumption 2.1 follows from the CLT for linear statistics and $G_\epsilon = G * \phi_\epsilon$ for some sufficiently nice mollifier ϕ . In the abstract context, we think of G_ϵ has a continuous approximation of G coming from a possibly different regularization procedure. To construct a multiplicative chaos measure out of the field $X_{N,\epsilon}$, it is clear that assumption 2.1 is necessary and that one also needs the existence of the GMC measure ν_T^γ .

Assumption 2.2 (Convergence of the GMC measure). *Let $\nu_{T,\epsilon}^\gamma(dx) = \exp\left(\gamma G_\epsilon(x) - \frac{\gamma^2}{2} \mathbb{E} [G_\epsilon(x)^2]\right) dx$ for any $\gamma, \epsilon > 0$. We suppose that, if $\gamma^2 q < 2d$, then for any $w \in L^1(\mathfrak{A})$, we have the convergence in $L^q(\mathbb{P})$*

$$\nu_{T,\epsilon}^\gamma(w) \Rightarrow \nu_T^\gamma(w) \quad \text{as } \epsilon \rightarrow 0. \quad (2.3)$$

As discussed in the proof of proposition 1.7 and the remark below, (2.3) basically follows from the asymptotics of the covariance kernel (2.2). Namely, we will require the following conditions.

Assumption 2.3 (Covariance kernel asymptotics). *Suppose that for all $(u, v) \in \mathfrak{A}^2$,*

$$T_{\epsilon,\delta}(u, v) \leq \log^+ (|u - v|^{-1} \wedge \epsilon^{-1} \wedge \delta^{-1}) + C, \quad (2.4)$$

and that for almost all $(u, v) \in \mathfrak{A}^2$,

$$T_{\epsilon,\delta}(u, v) \rightarrow T(u, v) \quad \text{as } \epsilon, \delta \rightarrow 0. \quad (2.5)$$

We also suppose the bound (2.4) is sharp, in the sense that if $\delta \geq \epsilon \geq 0$ and $|u - v| \leq \exp(-\delta^{-1})$, then

$$T_{\epsilon,\delta}(u, v) = \log \delta^{-1} + \underset{\delta \rightarrow 0}{O}(1). \quad (2.6)$$

Finally, in order to apply the *second moment method* considered by Webb in [68], we will also need to control some exponential moments of the field $(u, \epsilon) \mapsto \tilde{X}_{N,\epsilon}(u)$. The idea of [68] consists in proving that both in the weak regime (when we consider successive limits as $N \rightarrow \infty$ and $\epsilon \rightarrow 0$) and in the strong regime (when $\epsilon_N \rightarrow 0$ as $N \rightarrow \infty$), the limiting random measures coincide and have the same law as the GMC measure ν_T^γ coming directly from a log-correlated process G . In particular, we will need the following asymptotics.

Assumption 2.4 (Exponential moments asymptotics). *Let $q \in \mathbb{N}$. We suppose that there exists a sequence $\delta_N \rightarrow 0$ as $N \rightarrow \infty$ so that for any $\epsilon \in \Delta_q(\delta_N)$, (1.47), and for any $\mathbf{t} \in \mathbb{R}^q$, we have uniformly for all $\mathbf{u} \in \mathfrak{A}^q$,*

$$\log \mathbb{E} \left[\exp \left(\sum_{k=1}^q \tilde{X}_{N,\epsilon_k}^{t_k}(u_k) \right) \right] = \sum_{1 \leq k < j \leq q} t_k t_j T_{\epsilon_j, \epsilon_k}(u_j, u_k) + o(1). \quad (2.7)$$

Theorem 2.1. *Suppose that the assumptions 2.1 – 2.3 hold, as well as the assumption 2.4 with $q = 1, 2$. Let ϵ_N be any sequence which converges to 0 as $N \rightarrow \infty$ in such a way that $\epsilon_N \geq \delta_N$. If $\gamma^2 < d$, for any $w \in L^1(\mathbb{R})$, then the random variable $\mu_{N,\epsilon_N}^\gamma(w)$ converges in distribution as $N \rightarrow \infty$ to $\nu_T^\gamma(w)$. Moreover, if the assumption 2.4 is also satisfied for $q \in \mathbb{N}$ and $\gamma^2 q < 2d$, then*

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\mu_{N,\epsilon_N}^\gamma(w)^q \right] = \int \exp \left(\gamma^2 \sum_{1 \leq j < k \leq q} T(u_j, u_k) \right) \prod_{k=1}^q w(u_k) du_k. \quad (2.8)$$

As mentioned in the introduction, the condition $\gamma^2 q < 2d$ for the existence of the limiting moments (2.8) is sharp and is related to the convergence of certain Selberg integrals. The remainder of this section is devoted to the proof of theorem 2.1 and of an extension, proposition 2.5, in the case where there is a coupling of the fields $X_{N,\epsilon}(u)$ for different $N > 0$. This applies for instance to the sine process discussed in section 1.6. Note that to prove the existence of the measure μ^γ in the L^2 -phase and the convergence of moments, it is clear from our assumptions that one can use the same argument as in the proof of proposition 1.7. However, to identify that this measure has the same distribution as ν_T^γ , it is simpler to first establish that $\mu_{N,\epsilon}^\gamma$ converges in distribution in the weak regime and then to show that the strong and weak limits coincide; c.f. lemma 2.2 and lemma 2.4 respectively.

Lemma 2.2. *Let $w \in L^1(\mathfrak{A})$ and fix $\epsilon > 0$. For any $\gamma \geq 0$, the random variable $\mu_{N,\epsilon}^\gamma(w)$ converges in distribution as $N \rightarrow \infty$ to $\nu_{T,\epsilon}^\gamma(w)$.*

Proof. By assumption 2.1 and continuity of the exponential function, the finite dimensional distributions of the process $\xi_{N,\epsilon}(u) = \exp \left(\tilde{X}_{N,\epsilon}^\gamma(u) \right)$ converge to those of $\xi_\epsilon(u) = \exp \left(\gamma G_\epsilon(u) - \frac{\gamma^2}{2} \mathbb{E}(G_\epsilon(u)^2) \right)$ as $N \rightarrow \infty$. We also claim that $\xi_{N,\epsilon}(u)$ is tight in $L^1(\mathbb{R}^d, dw)$, so that, by Prokhorov's theorem, $\xi_{N,\epsilon} \Rightarrow \xi_\epsilon$ as $N \rightarrow \infty$. The tightness follows from the criteria established in [14] which show that it suffices that there exists a constant $C_\epsilon > 0$ so that

$$\mathbb{E} [|\xi_{N,\epsilon}(u)|] \leq C_\epsilon \quad \text{and} \quad \mathbb{E} [|\xi_{N,\epsilon}(u)|] \rightarrow \mathbb{E} [|\xi_\epsilon(u)|]. \quad (2.9)$$

Notice that, since G_ϵ is a Gaussian process $\mathbb{E} [|\xi_\epsilon(u)|] = 1$ and the estimates (2.9) follow directly from the assumption 2.4. Since the functional $\xi \rightarrow \int \xi(u)w(u)du$ is obviously continuous on $L^1(\mathbb{R}^d, dw)$, we conclude that as $N \rightarrow \infty$,

$$\mu_{N,\epsilon}^\gamma(w) = \int \xi_{N,\epsilon}(u)w(u)du \Rightarrow \nu_{T,\epsilon}^\gamma(w) = \int \xi_\epsilon(u)w(u)du \quad \square$$

Remark 2.3. *This proof relies on the fact that the sequence $(\xi_{N,\epsilon})_{N>0}$ is tight in $L^1(\mathbb{R}^d, dw)$ for any $\epsilon > 0$ and that the conditions (2.9) are easy to check. However, for the CUE or sine statistics (c.f. (1.5) and (1.39) respectively), using the specific form of the test function $\chi_u \star \phi_\epsilon$, it is also possible to verify that the criterion (4) of [40, Theorem 16.5] holds, which implies that $X_{N,\epsilon} \Rightarrow X_\epsilon$ as random elements of $C(\mathfrak{A} \rightarrow \mathbb{R})$. This provides an alternative way to prove lemma 2.2.*

Lemma 2.4. *Let q be an even integer such that $\gamma^2 q < 2d$. Then, for any $w \in L^1(\mathfrak{A})$,*

$$\lim_{\epsilon \rightarrow 0, N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\delta_N}^\gamma(w) - \mu_{N,\epsilon}^\gamma(w) \right|^q \right] = 0.$$

Proof. Suppose that $\delta_N \leq \epsilon$. For any $i \in [q]$, define

$$\boldsymbol{\epsilon}^i = \underbrace{(\epsilon, \dots, \epsilon)}_{i\#} \underbrace{(\delta_N, \dots, \delta_N)}_{(q-i)\#}$$

and let for any $\mathbf{u} \in \mathfrak{A}^q$,

$$\Theta_{\epsilon, N}^i(\mathbf{u}) = \sum_{1 \leq k < j \leq q} T_{\epsilon_j^i, \epsilon_k^i}(u_j, u_k).$$

By Fubini's theorem and the assumption 2.4 with $\mathbf{t} = (\gamma, \dots, \gamma)$, we obtain

$$\mathbb{E} \left[\left| \mu_{N, \delta_N}^\gamma(w) - \mu_{N, \epsilon}^\gamma(w) \right|^q \right] = \sum_{i=1}^q (-1)^i \binom{q}{i} \int_{\mathfrak{A}^q} \exp \left(\gamma^2 \Theta_{\epsilon, N}^i(\mathbf{u}) + o(1) \right) \prod_{k=1}^q w(u_k) du_k, \quad (2.10)$$

where the error term is uniform. Moreover, the condition (2.5) implies that for any $i \in [q]$ and for almost all $\mathbf{u} \in \mathfrak{A}^q$,

$$\lim_{\epsilon \rightarrow 0, N \rightarrow \infty} \Theta_{\epsilon, N}^i(\mathbf{u}) = \sum_{1 \leq k < j \leq n} T(u_j, u_k).$$

Finally, the condition (2.4) shows that for all $\mathbf{u} \in \mathfrak{A}^q$,

$$\exp \left(\gamma^2 \Theta_{\epsilon, N}^i(\mathbf{u}) \right) \leq C \prod_{1 \leq j < k \leq q} 1 \vee |u_j - u_k|^{-\gamma^2}. \quad (2.11)$$

Hence, since the RHS of (2.11) is locally integrable on $(\mathbb{R}^d)^q$ when $\gamma^2 q < 2d$, by the dominated convergence theorem, all the integrals on the RHS of formula (2.10) converge to the same finite value while taking the limit as $N \rightarrow \infty$ and then as $\epsilon \rightarrow 0$. Since $\sum_{i=1}^q (-1)^i \binom{q}{i} = 0$, this proves the claim. \square

Proof of theorem 2.1. By lemma 2.2, for any $\epsilon > 0$, we have $\mu_{N, \epsilon}^\gamma(w) \Rightarrow \nu_{T, \epsilon}^\gamma(w)$ as $N \rightarrow \infty$ and the assumption 2.2 guarantees that $\nu_{T, \epsilon}^\gamma(w) \Rightarrow \nu_T^\gamma(w)$ as $\epsilon \rightarrow 0$. In addition, by lemma 2.4 with $q = 2$, we have proved that, almost surely,

$$\mu_{N, \delta_N}^\gamma(w) - \mu_{N, \epsilon}^\gamma(w) \rightarrow 0$$

as $N \rightarrow \infty$ and then $\epsilon \rightarrow 0$. By [40, Theorem 4.28], these conditions imply that $\mu_{N, \delta_N}^\gamma(w) \Rightarrow \nu_T^\gamma(w)$ as $N \rightarrow \infty$. To complete the proof, it remains to establish convergence of the moments of the random variable $\mu_{N, \delta_N}^\gamma(w)$. We proceed like in the proof of lemma 2.4. By Fubini's theorem, for any $q \in \mathbb{N}$ and $\epsilon > 0$, we have

$$\mathbb{E} \left[\mu_{N, \epsilon}^\gamma(w)^q \right] = \int_{\mathbb{R}^q} \mathbb{E} \left[\exp \left(\sum_{k=1}^q \tilde{X}_{N, \epsilon}^\gamma(u_j) \right) \right] \prod_{k=1}^q w(u_k) du_k. \quad (2.12)$$

Then, the assumptions 2.3–2.4 imply that for almost all $\mathbf{u} \in \mathfrak{A}^q$,

$$\lim_{\substack{N \rightarrow \infty \\ \epsilon \rightarrow 0}} \mathbb{E} \left[\exp \left(\sum_{k=1}^q \tilde{X}_{N, \epsilon}^\gamma(u_j) \right) \right] = \exp \left(\gamma^2 \sum_{1 \leq j < k \leq q} \mathbb{E} [G(u_j) G(u_k)] \right) \quad (2.13)$$

in both the weak and strong regime (as long as $\epsilon_N \geq \delta_N$). Moreover, the condition (2.4) guarantees that for all $\mathbf{u} \in \mathbb{R}^q$,

$$\mathbb{E} \left[\exp \left(\sum_{k=1}^q \tilde{X}_{N, \epsilon}^\gamma(u_j) \right) \right] \leq C \prod_{1 \leq j < k \leq q} 1 \vee |u_j - u_k|^{-\gamma^2}.$$

Hence, if $\gamma^2 q < 2$, formula (2.8) follows directly from (2.13) and the dominated convergence theorem. \square

Note that in the context of theorem 2.1 we did not require that the fields $X_{N,\epsilon}, X_{N+1,\epsilon}, \dots$ are defined on the same probability space. However, if we have such a good coupling like for the sine process, then we can upgrade the topology of convergence in theorem 2.1 by using the following result instead of lemma 2.2

Proposition 2.5. *Fix $\epsilon > 0$ and for any $N \in \mathbb{N}$, let $X_{N,\epsilon}(u) = \sum_{\lambda \in \Lambda_1} \chi_u * \phi_\epsilon(\lambda/N)$ where Λ_1 is the sine process and consider the measure $\mu_{N,\epsilon}^\gamma(du) = \exp\left(\tilde{X}_{N,\epsilon}^\gamma(u)\right) du$, c.f. formulae (1.39)–(1.42). For any $w \in L^1(\mathbb{R})$ and for any $q \geq 1$, the random variable $\mu_{N,\epsilon}^\gamma(w)$ converges in $L^q(\mathbb{P})$ as $N \rightarrow \infty$ to a limit $\mu_\epsilon^\gamma(w)$ whose law is the same as $\nu_{T,\epsilon}^\gamma(w)$.*

Proof. To keep the notation simple, we will prove the proposition when $q = 2$ which is the most interesting case. It is straightforward to generalize the argument to any even q . Like in the proof of theorem 1.9 (c.f. formula (1.48)), it is easy to check that the asymptotics of proposition 1.11 implies that for any $\eta, \eta' \in \{0, 1\}$,

$$\log \mathbb{E} \left[\exp \left(\tilde{X}_{N+\eta,\epsilon}^\gamma(u) + \tilde{X}_{N+\eta',\epsilon}^\gamma(v) \right) \right] = \gamma^2 T_{\epsilon,\epsilon}(u, v) + O_{\eta,\eta'} \left(e^{-(\log N)^\beta} \right)$$

uniformly for all $u, v \in \mathbb{R}$, where $T_{\epsilon,\epsilon}(u, v) = \mathbb{E}[G_{\phi,\epsilon}(u)G_{\phi,\epsilon}(v)]$. Expanding the square, this implies that

$$\mathbb{E} \left[\left| \mu_{N,\epsilon}^\gamma(w) - \mu_{N+1,\epsilon}^\gamma(w) \right|^2 \right] = \int_{\mathbb{R}} e^{\gamma^2 T_{\epsilon,\epsilon}(u,v)} \left\{ \sum_{\eta,\eta' \in \{0,1\}} (-1)^{\eta+\eta'} \exp O_{\eta,\eta'} \left(e^{-(\log N)^\beta} \right) \right\} w(u)w(v) du dv. \quad (2.14)$$

Note that the leading terms on the RHS of formula (2.14) cancel and the error terms are uniform. Moreover, by (2.4), $T_{\epsilon,\epsilon}(u, v) \leq \log^+(\epsilon^{-1}) + C$ on \mathbb{R}^2 and since $w \in L^1(\mathbb{R})$, we obtain for any $\epsilon > 0$,

$$\mathbb{E} \left[\left| \mu_{N,\epsilon}^\gamma(w) - \mu_{N+1,\epsilon}^\gamma(w) \right|^2 \right] = O_{w,\epsilon} \left(e^{-(\log N)^\beta} \right).$$

Thus, since the parameter $\beta > 1$, by the triangle inequality, this shows that $(\mu_{N,\epsilon}^\gamma(w))_{N>0}$ is a Cauchy sequence in $L^2(\mathbb{P})$. Let us denote by $\mu_\epsilon^\gamma(w)$ its limit. To complete the proof, we will use lemma 2.2 to prove that for any $\epsilon > 0$, $\mu_{N,\epsilon}^\gamma(w) \stackrel{d}{=} \nu_{T,\epsilon}^\gamma(w)$. For any $M > 0$, we let $w_M(u) = w(u)\mathbb{1}_{\{|u| \leq M\}}$ and $w_M^*(u) = w(u)\mathbb{1}_{\{|u| > M\}}$. On the one hand, since

$$\limsup_{N \rightarrow \infty} \mathbb{E} \left[\mu_{N,\epsilon}^\gamma(w_M^*)^2 \right] \leq C\epsilon^{-1} \int_{|u| > M} w(u) du,$$

the random variable $\mu_\epsilon^\gamma(w_M)$ converges to $\mu_\epsilon^\gamma(w)$ in $L^2(\mathbb{P})$ as $M \rightarrow \infty$. In addition, a similar estimate shows that for the regularized GMC measure: $\nu_{T,\epsilon}^\gamma(w_M) \Rightarrow \nu_{T,\epsilon}^\gamma(w)$ as $M \rightarrow \infty$.

On the other hand, by lemma 2.2, $\mu_\epsilon^\gamma(w_M) \stackrel{d}{=} \nu_{T,\epsilon}^\gamma(w_M)$ for any $M > 0$. Hence, we conclude that $\mu_\epsilon^\gamma(w) \stackrel{d}{=} \nu_{T,\epsilon}^\gamma(w)$. \square

2.2 Convergence of the multiplicative chaos measure in the L^1 -phase

We work in the same context as in the previous section and, under slightly stronger assumptions than that of theorem 2.1, we will establish the convergence of the random measure $\mu_{N,\epsilon}^\gamma(du) = \exp(\tilde{X}_{N,\epsilon}^\gamma(u)) du$ on $\mathfrak{A} \subset \mathbb{R}^d$ when $0 < \gamma < \sqrt{2d}$. By analogy with GMC theory, this condition

should be sharp. That is, in the strong regime, with the normalization (2.1), we expect that $\mu_{N,\epsilon}^\gamma(du) \rightarrow 0$ for any $\gamma \geq \sqrt{2d}$. Before stating our main result, we need to introduce some further notation. For any $u \in \mathbb{R}$, let $Z^N(u)$ be the random variable taking value in $\mathbb{R}^{\mathbb{N}}$ (defined with respect to the law \mathbb{P} of the point process Λ_N) given by

$$Z_k^N(u) = \begin{cases} X_{N,e^{-k}}(u) & \text{if } k \leq \log \delta_N^{-1} \\ 0 & \text{else} \end{cases}. \quad (2.15)$$

For any $u \in \mathfrak{A}$, we denote by \mathbb{G}_u the Gaussian measure on $\mathfrak{A}^{\mathbb{N}}$ defined by $\mathbf{E}_{\mathbb{G}_u}[Z_k] = \gamma T_{0,e^{-k}}(u, u)$ for all $k \in \mathbb{N}$, and whose covariance structure is given by

$$\langle Z_k; Z_j \rangle_{\mathbb{G}_u} = T_{e^{-j}, e^{-k}}(u, u).$$

We also let $\mathbb{G}_{u,v}$ be the Gaussian measure on $\mathfrak{A}^{\mathbb{N}} \times \mathfrak{A}^{\mathbb{N}}$ defined as follows. If $(Z(u), Z(v)) \sim \mathbb{G}_{u,v}$, then we have

$$\mathbf{E}_{\mathbb{G}_{u,v}}[Z_k(x)] = \gamma(T_{0,e^{-k}}(x, x) + T_{0,e^{-k}}(u, v)), \quad \forall x \in \{u, v\},$$

and the covariance structure:

$$\langle Z_k(x); Z_j(y) \rangle_{\mathbb{G}_{u,v}} = T_{e^{-k}, e^{-j}}(x, y), \quad \forall x, y \in \{u, v\}.$$

For any $u \in \mathfrak{A}$, we denote by $\mathbb{P}_{N,\epsilon}^u$ be the probability measure with Radon-Nykodym derivative proportional to $\exp(\tilde{X}_{N,\epsilon}(u))$ with respect to the probability measure \mathbb{P} . Finally, for any $(u, v) \in \mathfrak{A}^2$ such that $u \neq v$, we let $\mathbb{P}_{N,\epsilon}^{u,v}$ be the probability measure with Radon-Nykodym derivative proportional to $\exp(\tilde{X}_{N,\epsilon}(u) + \tilde{X}_{N,\epsilon}(v))$ and, in the mixed regime, we let $\tilde{\mathbb{P}}_{N,\epsilon}^{u,v}$ be the probability measure with Radon-Nykodym derivative proportional to $\exp(\tilde{X}_{N,\epsilon_N}(u) + \tilde{X}_{N,\epsilon}(v))$ with respect to \mathbb{P} .

The point of these definitions is that we make the following assumption to prove our main result.

Assumption 2.5 (Weak convergence under the biased measures). *For any $u \in \mathfrak{A}$, we have*

$$\mathbf{Law}_{\mathbb{P}_{N,\epsilon_N}^u}(Z^N(u)) \Rightarrow \mathbb{G}_u \quad \text{as } N \rightarrow \infty, \quad (2.16)$$

$$\mathbf{Law}_{\mathbb{P}_{N,\epsilon}^u}(Z^N(u)) \Rightarrow \mathbb{G}_u \quad \text{as } N \rightarrow \infty \text{ and then } \epsilon \rightarrow 0. \quad (2.17)$$

For any $(u, v) \in \mathfrak{A}^2$ such that $u \neq v$, we have

$$\mathbf{Law}_{\mathbb{P}_{N,\epsilon_N}^{u,v}}(Z^N(u), Z^N(v)) \Rightarrow \mathbb{G}_{u,v} \quad \text{as } N \rightarrow \infty, \quad (2.18)$$

$$\mathbf{Law}_{\mathbb{P}_{N,\epsilon}^{u,v}}(Z^N(u), Z^N(v)) \Rightarrow \mathbb{G}_{u,v} \quad \text{as } N \rightarrow \infty \text{ and then } \epsilon \rightarrow 0. \quad (2.19)$$

$$\mathbf{Law}_{\tilde{\mathbb{P}}_{N,\epsilon}^{u,v}}(Z^N(u), Z^N(v)) \Rightarrow \mathbb{G}_{u,v} \quad \text{as } N \rightarrow \infty \text{ and then } \epsilon \rightarrow 0. \quad (2.20)$$

Theorem 2.6. *Suppose that the assumptions 2.1 – 2.5 hold for $q \leq 3$ in 2.4. Let $w \in L^1(\mathfrak{A})$ and ϵ_N be a sequence which converges to 0 as $N \rightarrow \infty$ in such a way that $\epsilon_N \geq \delta_N$. Then, if $\gamma^2 < 2d$, the random variable $\mu_{N,\epsilon_N}^\gamma(w)$ converges in distribution to $\nu_T^\gamma(w)$ as $N \rightarrow \infty$.*

The proof of theorem 2.6 follows the elementary argument introduced [4] to prove convergence of the GMC measure in the L^1 -phase. In fact, the asymptotics (2.7) are so strong that Berestycki's method can be applied to the field $u \mapsto X_{N,\epsilon}(u)$ only modulo a few technical details. The main idea stems from the fact that the measure $\mu_{N,\epsilon}^\gamma$ is supported on the so-called γ -thick points which satisfies:

$$\lim_{N \rightarrow \infty} \frac{X_{N,\epsilon_N}(u)}{\log \epsilon_N^{-1}} = \gamma.$$

More specifically, we will proceed to show that, if the parameter $\alpha > \gamma$, then the mass

$$\mu_{N,\epsilon_N}^\gamma \left(u \in \mathfrak{A} : \frac{X_{N,\epsilon}(u)}{\log \epsilon} > \alpha \text{ for all } \epsilon \in \{e^{-k} : L \leq k \leq \log \delta_N^{-1}\} \right)$$

converges to 0 in $L^1(\mathbb{P})$ as $N \rightarrow \infty$ and then $L \rightarrow \infty$, and to show that the random measures $\mu_{N,\epsilon_N}^\gamma$ restricted to the good set $\left\{ |u| \in \mathfrak{A} : \frac{X_{N,\epsilon}(u)}{\log \epsilon} \leq \alpha \text{ for all } \epsilon \in \{e^{-k} : L \leq k \leq \log \delta_N^{-1}\} \right\}$ form a Cauchy sequence in $L^2(\mathbb{P})$ when α is sufficiently close to γ and $\gamma < \sqrt{2d}$; c.f. lemmas 2.7 and proposition 2.8 respectively. Moreover, like in section 2.1, we will rely on the fact that the weak and strong limits coincide to identify that the law of the random variable μ^γ is the same law as the GMC measure ν_T^γ . Theorem 2.6 will be proved after we establish the preparatory lemmas. Then, we will show that the assumption 2.5 automatically holds if the assumption 2.4 is true for all $q \in \mathbb{N}$; c.f. proposition 2.9. Finally, in the end, we will give a quick proof of theorem 1.9.

For any $\alpha > 0$ and $L > 0$, we define the following events:

$$A_L^\alpha(Z) = \{Z_k \leq \alpha k, \forall k \geq L\} \quad \text{and} \quad A_L^{\alpha*}(Z) = \{Z_k > \alpha k, \forall k \geq L\}. \quad (2.21)$$

Lemma 2.7. *Under the assumptions of theorem 2.6, if $\alpha > \gamma$, then we have*

$$\lim_{L \rightarrow \infty} \lim_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\epsilon_N}^\gamma \left(w \mathbb{1}_{A_L^{\alpha*}(Z^N)} \right) \right| \right] = 0, \quad (2.22)$$

$$\lim_{L \rightarrow \infty} \lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\epsilon}^\gamma \left(w \mathbb{1}_{A_L^{\alpha*}(Z^N)} \right) \right| \right] = 0. \quad (2.23)$$

Proof. We shall only prove formula (2.22) since the proof of (2.23) is exactly the same as the asymptotics (2.7) which holds both in the weak and the strong regime (with $\epsilon_N = \delta_N$). Moreover, without loss of generality, we assume that the function $w \geq 0$. By definition of the probability measure $\mathbb{P}_{N,\epsilon}^u$, we may rewrite for any $N > 0$ and $\epsilon > 0$,

$$\mathbb{E} \left[\mu_{N,\epsilon}^\gamma \left(w \mathbb{1}_{A_L^{\alpha*}(Z^N)} \right) \right] = \int_{\mathbb{R}^d} \mathbb{P}_{N,\epsilon}^u [A_L^{\alpha*}(Z^N(u))] \mathbb{E} \left[e^{\tilde{X}_{N,\epsilon}(u)} \right] w(u) du.$$

First, the assumption 2.4 implies that $\mathbb{E} \left[e^{\tilde{X}_{N,\delta_N}(u)} \right] \rightarrow 1$ as $N \rightarrow \infty$ uniformly for all $u \in \mathfrak{A}$. Second, by (2.16) and the Portmanteau theorem,

$$\lim_{N \rightarrow \infty} \mathbb{P}_{N,\delta_N}^u [A_L^{\alpha*}(Z^N(u))] = \mathbb{G}_u [A_L^{\alpha*}(Z)].$$

Note that since \mathbb{G}_u is a Gaussian measure on $\mathbb{R}^{\mathbb{N}}$, $A_L^{\alpha*}$ is a set of continuity for \mathbb{G}_u . Hence, since $w \in L^1(\mathfrak{A})$, by the dominated convergence theorem, we obtain

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\mu_{N,\epsilon}^\gamma \left(w \mathbb{1}_{A_L^{\alpha*}(Z^N)} \right) \right] = \int_{\mathbb{R}^d} \mathbb{G}_u [A_L^{\alpha*}(Z)] w(u) du.$$

Thus, in order to prove that (2.22), it suffices to show that the probability $\mathbb{G}_u [A_L^{\alpha*}(Z)] \rightarrow 0$ as $L \rightarrow \infty$. This is a standard Gaussian tail-bound computation. Note that the condition (2.6) implies that

$$\mathbf{E}_{\mathbb{G}_u} [Z_k] \leq \gamma k + C \quad \text{and} \quad \langle Z_k; Z_k \rangle_{\mathbb{G}_u} = \mathbf{Var}_{\mathbb{G}_u} [Z_k] = k + O(1). \quad k \rightarrow \infty$$

Thus, since $\alpha > \gamma$, if the parameter L is sufficiently large, we can assume that for all $k \geq L$,

$$\alpha k - \mathbf{E}_{\mathbb{G}_u} [Z_k] \geq \sqrt{3}(\alpha - \gamma)k/2 \quad \text{and} \quad \mathbf{Var}_{\mathbb{G}_u} [Z_k] \leq 3k/2.$$

Then, a union bound and a Gaussian tail-bound show that for any $u \in \mathfrak{A}$,

$$\begin{aligned}\mathbb{G}_u[A_L^{\alpha*}(Z)] &\leq \sum_{k \geq L} \mathbb{G}_u[Z_k > \alpha k] \leq \sum_{k \geq L} \exp\left(-\frac{(\alpha k - \mathbf{E}_{\mathbb{G}_u}[Z_k])^2}{2\mathbf{Var}_{\mathbb{G}_u}[Z_k]}\right) \\ &\leq \sum_{k \geq L} \exp\left(-\frac{(\alpha - \gamma)^2}{4}k\right).\end{aligned}$$

Hence, $\mathbb{G}_u[A_L^{\alpha*}(Z)] \rightarrow 0$ as $L \rightarrow \infty$ and this completes the proof. \square

Proposition 2.8. *Let $\gamma > 0$ and $\alpha = \gamma + ((2d\gamma^{-1} - \gamma) \wedge \gamma)/2$ (in particular, we can choose the parameter $\gamma < \alpha < 3\gamma/2$ only if $\gamma^2 < 2d$). Then, under the assumptions of theorem 2.6, for any $L > 0$, we have*

$$\lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N, \delta_N}^\gamma(w \mathbb{1}_{A_L^\alpha(Z^N)}) - \mu_{N, \epsilon}^\gamma(w \mathbb{1}_{A_L^\alpha(Z^N)}) \right|^2 \right] = 0.$$

Proof. As in lemma 2.7, we let $\epsilon_N = \delta_N$. For any $(u, v) \in \mathfrak{A}^2$, $\epsilon > 0$ and $N > 0$, let us define

$$\mathscr{W}_{N, \epsilon}^{u, v} = \mathbb{E} \left[\exp(\tilde{X}_{N, \epsilon}(u) + \tilde{X}_{N, \epsilon}(v)) \right] \quad \text{and} \quad \widetilde{\mathscr{W}}_{N, \epsilon}^{u, v} = \mathbb{E} \left[\exp(\tilde{X}_{N, \delta_N}(u) + \tilde{X}_{N, \epsilon}(v)) \right].$$

We also let

$$\mathcal{I}_{N, \epsilon} = \iint_{\mathbb{R}^2} \mathbb{P}_{N, \epsilon}^{u, v}[A_L^\alpha(Z^N(u)), A_L^\alpha(Z^N(v))] \mathscr{W}_{N, \epsilon}^{u, v} w(u) w(v) du dv$$

and

$$\tilde{\mathcal{I}}_{N, \epsilon} = \iint_{\mathbb{R}^2} \tilde{\mathbb{P}}_{N, \epsilon}^{u, v}[A_L^\alpha(Z^N(u)), A_L^\alpha(Z^N(v))] \widetilde{\mathscr{W}}_{N, \epsilon}^{u, v} w(u) w(v) du dv.$$

Like in the proof of proposition 2.4, we may expand

$$\mathbb{E} \left[\left| \mu_{N, \delta_N}^\gamma(w \mathbb{1}_{A_L(Z^N)}) - \mu_{N, \epsilon}^\gamma(w \mathbb{1}_{A_L(Z^N)}) \right|^2 \right] = \mathcal{I}_{N, \delta_N} + \mathcal{I}_{N, \epsilon} - 2\tilde{\mathcal{I}}_{N, \epsilon} \quad (2.24)$$

and we would like to prove that all the terms converge to the same limit when $N \rightarrow \infty$ and then $\epsilon \rightarrow 0$. We will focus on computing the limit of the integral $\mathcal{I}_{N, \delta_N}$ as $N \rightarrow \infty$. On the one hand, the assumption (2.18) implies that for all $u, v \in \mathfrak{A}$ so that $u \neq v$,

$$\lim_{N \rightarrow \infty} \mathbb{P}_{N, \delta_N}^{u, v}[A_L^\alpha(Z^N(u)), A_L^\alpha(Z^N(v))] = \mathbb{G}_{u, v}[A_L^\alpha(Z(u)), A_L^\alpha(Z(v))]. \quad (2.25)$$

On the other hand, the assumption 2.4 implies that for all $(u, v) \in \mathfrak{A}^2$,

$$\mathscr{W}_{N, \delta_N}^{u, v} \sim \exp(\gamma^2 T_{\delta_N, \delta_N}(u, v)) \quad \text{as } N \rightarrow \infty. \quad (2.26)$$

Moreover, by (2.4),

$$\mathscr{W}_{N, \delta_N}^{u, v} \ll (|u - v| \vee \delta_N)^{-\gamma^2}, \quad (2.27)$$

and we expect that, in the regime where $\gamma > 1$, the limit of the integral $\mathcal{I}_{N, \delta_N}$ is finite only if the probability the probability $\mathbb{G}_{u, v}[A_L^\alpha(Z(u)), A_L^\alpha(Z(v))]$ converge to 0 sufficiently fast as $|u - v| \rightarrow 0$. In order to prove this, we use that $A_L^\alpha(Z) \subset \{Z_\zeta \leq \alpha\zeta\}$ where

$$\zeta = \lfloor (\log |u - v|^{-1} \vee L) \wedge (\log \delta_N^{-1}) \rfloor,$$

so that

$$\mathbb{P}_{N, \delta_N}^{u, v}[A_L^\alpha(Z^N(u)), A_L^\alpha(Z^N(v))] \leq \mathbb{P}_{N, \delta_N}^{u, v}[Z_\zeta^N \leq \alpha\zeta].$$

By Markov's inequality and the asymptotics (2.7), this implies that for any $t < 0$,

$$\begin{aligned}\mathbb{P}_{N,\delta_N}^{u,v} [Z_\zeta^N \leq \alpha\zeta] &\leq \frac{\mathbb{E} [\exp (\gamma X_{N,\delta_N}(u) + \gamma X_{N,\delta_N}(v) - t X_{N,e^{-\zeta}}(u))]}{\mathbb{E} [\exp (\gamma X_{N,\delta_N}(u) + \gamma X_{N,\delta_N}(v))]} e^{\alpha t \zeta} \\ &= \exp \left(\alpha t \zeta - \gamma t (T_{\delta_N,e^{-\zeta}}(u,u) + T_{\delta_N,e^{-\zeta}}(u,v)) + \frac{t^2}{2} T_{e^{-\zeta},e^{-\zeta}}(u,u) + o(1) \right)_{N \rightarrow \infty}.\end{aligned}$$

Then, choosing $t = t_*(N, \gamma, \alpha, u, v) := (\alpha\zeta - \gamma(T_{\delta_N,e^{-\zeta}}(u,u) + T_{\delta_N,e^{-\zeta}}(u,v))) / T_{e^{-\zeta},e^{-\zeta}}(u,u)$, we obtain

$$\mathbb{P}_{N,\delta_N}^{u,v} [Z_\zeta^N \leq \alpha\zeta] \ll \exp \left(-\frac{t_*^2}{2} T_{e^{-\zeta},e^{-\zeta}}(u,u) \right). \quad (2.28)$$

Note that by definition of ζ , the assumption (2.6) shows that $t_* = (\alpha - 2\gamma) + o(1)$. In particular, if the parameter L is sufficiently large and $\alpha < 2\gamma$, the parameter $t_* < 0$ so that the bound (2.28) holds and we have for all $u, v \in \mathbb{R}$,

$$\mathbb{P}_{N,\delta_N}^{u,v} [Z_\zeta^N \leq \alpha\zeta] \ll \exp \left(-\frac{(\alpha - 2\gamma)^2}{2} \zeta \right).$$

This shows that

$$\mathbb{P}_{N,\delta_N}^{u,v} [A_L^\alpha(Z^N(u)), A_L^\alpha(Z^N(v))] \ll (|u - v| \vee \delta_N)^{\frac{(\alpha - 2\gamma)^2}{2}}.$$

This estimate combined with (2.27) shows that when N is sufficiently large,

$$\mathbb{P}_{N,\epsilon}^{u,v} [A_L^\alpha(Z^N(u)), A_L^\alpha(Z^N(v))] \mathscr{W}_{N,\delta_N}^{u,v} \ll (|u - v| \vee \delta_N)^{\frac{(\alpha - 2\gamma)^2}{2} - \gamma^2} \vee 1. \quad (2.29)$$

Thus, taking $\alpha = \gamma + \varepsilon/2\gamma$ for some $\varepsilon > 0$ sufficiently small, we see that $\frac{(\alpha - 2\gamma)^2}{2} - \gamma^2 > -\frac{\gamma^2 + \varepsilon}{2}$. In particular, as long as $\gamma^2 < 2d$, the RHS of (2.29) is locally integrable on $\mathbb{R}^{\frac{d}{2}} \times \mathbb{R}^d$. Hence, by the dominated convergence theorem, formulae (2.25) and (2.26), we conclude that $\lim_{N \rightarrow \infty} \mathcal{I}_{N,\delta_N} = \mathcal{I}_\infty^{\gamma,L}$ where

$$\mathcal{I}_\infty^{\gamma,L} := \iint_{\mathbb{R}^2} \mathbb{G}_{u,v} [A_L^\alpha(Z(u)), A_L^\alpha(Z(v))] e^{\gamma^2 T(u,v)} w(u) w(v) du dv < \infty.$$

Using the hypothesis (2.19) and (2.20), we can apply the same argument in the weak and mixed regimes as well and obtain

$$\lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \mathcal{I}_{N,\epsilon} = \lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \tilde{\mathcal{I}}_{N,\epsilon} = \mathcal{I}_\infty^{\gamma,L}.$$

In the end, combining these limits with formula (2.24), we obtain the claim. \square

Proof of theorem 2.6. Recall that by lemma 2.2, for any $\epsilon > 0$, $\mu_{N,\epsilon}^\gamma(w) \Rightarrow \nu_{T,\epsilon}^\gamma(w)$ as $N \rightarrow \infty$ and, according to the assumption 2.2, if $\gamma^2 q < 2d$, the random variable $\nu_{T,\epsilon}^\gamma(w) \Rightarrow \nu_T^\gamma(w)$ as $\epsilon \rightarrow 0$. Therefore, by [40, Theorem 4.28], it suffices to prove that almost surely,

$$\limsup_{\epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\delta_N}^\gamma(w) - \mu_{N,\epsilon}^\gamma(w) \right| \right] = 0 \quad (2.30)$$

as $N \rightarrow \infty$ and then $\epsilon \rightarrow 0$. By Cauchy-Schwarz and the triangle inequality, we have for any $L, \alpha > 0$ and $\epsilon > 0$,

$$\begin{aligned}\mathbb{E} \left[\left| \mu_{N,\delta_N}^\gamma(w) - \mu_{N,\epsilon}^\gamma(w) \right| \right] &\leq \mathbb{E} \left[\left| \mu_{N,\delta_N}^\gamma \left(w \mathbb{1}_{A_L^\alpha(Z^N)} \right) - \mu_{N,\epsilon}^\gamma \left(w \mathbb{1}_{A_L^\alpha(Z^N)} \right) \right| \right] \\ &\quad + \mathbb{E} \left[\left| \mu_{N,\delta_N}^\gamma \left(w \mathbb{1}_{A_L^{\alpha^*}(Z^N)} \right) \right| \right] + \mathbb{E} \left[\left| \mu_{N,\epsilon}^\gamma \left(w \mathbb{1}_{A_L^{\alpha^*}(Z^N)} \right) \right| \right].\end{aligned}$$

Then, by proposition 2.8, if $\gamma^2 < 2d$, choosing the parameter α appropriately, we obtain

$$\begin{aligned} \limsup_{\epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N, \delta_N}^\gamma(w) - \mu_{N, \epsilon}^\gamma(w) \right| \right] &\leq \limsup_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N, \delta_N}^\gamma \left(w \mathbf{1}_{A_L^{\alpha*}(Z^N)} \right) \right| \right] \\ &+ \limsup_{\epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N, \epsilon}^\gamma \left(w \mathbf{1}_{A_L^{\alpha*}(Z^N)} \right) \right| \right]. \end{aligned} \quad (2.31)$$

Since the LHS of (2.31) does not depend on the parameter $L > 0$, we can further take the limit as $L \rightarrow \infty$. By lemma 2.7, the RHS of (2.31) converge to 0 and we conclude that (2.30) holds. \square

In principle, it is possible to verify the assumption 2.5 without knowing all the asymptotics of all exponential moments of the random field $X_{N, \epsilon}(u)$. However, when it is the case, the assumption 2.5 follows from a routine computation. In particular, as a consequence of the asymptotics obtained in section 2.3 and section 3, this applies to the CUE and sine process.

Proposition 2.9. *Suppose that the assumptions 2.1, 2.3 and 2.4 hold for all $q \in \mathbb{N}$. Then the assumption 2.5 is satisfied. In particular, if the condition 2.2 holds as well, we can apply theorem 2.6.*

Proof. it is easy to see that the proofs of (2.16)–(2.20) all follow the same argument. Therefore, we will focus on proving (2.18). In particular, fix $u, v \in \mathfrak{A}$ and recall the definition of the Gaussian measure $\mathbb{G}_{u, v}$. The assumptions 2.1 and 2.4 imply that for any $\mathbf{t} \in \mathbb{R}^q$ and $\mathbf{u} \in \{u, v\}^q$, we have

$$\begin{aligned} \mathbf{E}_{\mathbb{P}_{N, \delta_N}^{u, v}} \left[\exp \left(\sum_{k=1}^n t_k X_{N, e^{-k}}(u_k) \right) \right] &= \frac{\mathbb{E} \left[\exp \left(\tilde{X}_{N, \delta_N}^\gamma(u) + \tilde{X}_{N, \delta_N}^\gamma(v) + \sum_{k=1}^n \tilde{X}_{N, e^{-k}}^{t_k}(u_k) + \frac{t_k^2}{2} T_{e^{-k}, e^{-k}}(u_k, u_k) \right) \right]}{\mathbb{E} \left[\exp \left(\tilde{X}_{N, \delta_N}^\gamma(u) + \tilde{X}_{N, \delta_N}^\gamma(v) \right) \right]} \\ &= \exp \left(\sum_{k=1}^n t_k \mathbf{E}_{\mathbb{G}_{u, v}}[Z_k(u_k)] + \frac{1}{2} \sum_{k, j=1}^n t_k t_j \langle Z_k(u_k); Z_j(u_j) \rangle_{\mathbb{G}_{u, v}} + o(1) \right)_{N \rightarrow \infty}. \end{aligned} \quad (2.32)$$

This proves that under the biased measure $\mathbb{P}_{N, \delta_N}^{u, v}$, the finite dimensional law of the random process $(Z^N(u), Z^N(v))$ converges as $N \rightarrow \infty$ to $\mathbb{G}_{u, v}$. Hence, by Prokhorov's theorem, [40, Theorem 16.3], to prove convergence in distribution it remains to show tightness in $\mathbb{R}^N \times \mathbb{R}^N$ (which is a complete metric space). Let $\sigma > 0$, $c > 0$, and define

$$\mathfrak{V} = \bigcap_{k \in \mathbb{N}} [-\sigma k - c, \sigma k + c].$$

By Tychonoff's theorem, the set \mathfrak{V} is compact in the product topology of $\mathbb{R}^{\mathbb{N}}$. So, to establish tightness, it suffices to show that for any $\epsilon > 0$, if the parameters σ and c are sufficiently large, then

$$\limsup_{N \rightarrow \infty} \left(\mathbb{P}_{N, \delta_N}^{u, v} [Z^N(u) \notin \mathfrak{V}] \vee \mathbb{P}_{N, \delta_N}^{u, v} [Z^N(v) \notin \mathfrak{V}] \right) \leq \epsilon. \quad (2.33)$$

Note that by symmetry, both probabilities on the LHS of (2.33) have the same value. If $n_N = \lfloor \log \delta_N^{-1} \rfloor$, a union bound shows that

$$\mathbb{P}_{N, \delta_N}^{u, v} [Z^N(u) \notin \mathfrak{V}] \leq \sum_{k=1}^{n_N} \mathbb{P}_{N, \delta_N}^{u, v} [|X_{N, e^{-k}}(u)| \geq \sigma k + c].$$

Moreover, by Markov's inequality, we have for any $t > 0$

$$\mathbb{P}_{N, \delta_N}^{u, v} [|X_{N, e^{-k}}(u)| \geq \sigma k + c] \leq e^{-t(\sigma k + c)} \left(\mathbf{E}_{\mathbb{P}_{N, \delta_N}^{u, v}} [\exp(t X_{N, e^{-k}}(u))] + \mathbf{E}_{\mathbb{P}_{N, \delta_N}^{u, v}} [\exp(-t X_{N, e^{-k}}(u))] \right)$$

By formula (2.32), for any $t \in \mathbb{R}$,

$$\mathbf{E}_{\mathbb{P}_{N,\delta_N}^{u,v}} [\exp(tX_{N,e^{-k}}(u))] = \exp\left(t\gamma(T_{0,e^{-k}}(u,u) + T_{0,e^{-k}}(u,v)) + \frac{t^2}{2}T_{e^{-k},e^{-k}}(u,u) + o(1)_{N \rightarrow \infty}\right)$$

and, using the estimate (2.4), we obtain

$$\mathbf{E}_{\mathbb{P}_{N,\delta_N}^{u,v}} [\exp(tX_{N,e^{-k}}(u))] \ll e^{k(2\gamma|t|+t^2/2)}.$$

Choosing for instance $t = 2\gamma$, this implies that

$$\limsup_{N \rightarrow \infty} \mathbb{P}_{N,\delta_N}^{u,v} [Z^N(u) \notin \mathfrak{Y}] \ll e^{-2\gamma c} \sum_{k=1}^{\infty} e^{-2\gamma k(\sigma-3\gamma)}.$$

Hence, if $\sigma > 3\gamma$ and $e^{-2\gamma c}$ is sufficiently small compared to ε , we obtain the bound (2.33). \square

Let us finish this section by completing the proof of the main result for the sine process, theorem 1.9. Note that in this case, we can use the coupling (1.39) and proposition 2.5, to show that the random measure $\mu_{N,\epsilon}^\gamma(du) = \exp(\tilde{X}_{N,\epsilon}^\gamma(u)) du$ converges in $L^1(\mathbb{P})$ to a GMC measure

Proof of theorem 1.9. According to the proof of theorem 1.10, the linear statistics (1.39) of the sine process satisfy the assumptions 2.1, 2.3 and 2.4 for all $q \in \mathbb{N}$. Then, by proposition 2.9, the assumption 2.5 holds too and we deduce the estimate (2.30). By the remark 1.8, if $\gamma < \sqrt{2d}$, the random variable $\mu_\epsilon^\gamma(w)$ constructed in proposition 2.5 converges in $L^1(\mathbb{P})$ as $\epsilon \rightarrow 0$ to a random variable $\mu^\gamma(w)$ which has the same law as $\nu_Q^\gamma(w)$. In other words, we have

$$\limsup_{\epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\epsilon}^\gamma(w) - \mu^\gamma(w) \right| \right] = 0.$$

Combining this limit with (2.30), by the triangle inequality, we conclude that for any $w \in L^1(\mathbb{R})$, we have

$$\limsup_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\delta_N}^\gamma(w) - \mu^\gamma(w) \right| \right] = 0. \quad \square$$

Remark 2.10. Let $\mathfrak{A} \subset \mathbb{R}$ be a compact interval, $0 < \alpha < 1$, and let $\delta_N = (\log N)^2 N^{-\alpha}$. According to proposition 3.3 and corollary 2.14, the field $X_{N,\epsilon}(u)$ coming from regularized counting statistics of the mesoscopic CUE also satisfies the assumptions 2.1 – 2.4 for all $q \in \mathbb{N}$ and any Schwartz-class mollifier ϕ . Hence, by theorem 2.1 and proposition 2.9, this yields the full proof of theorem 1.2 stated in the introduction.

2.3 Asymptotics of the covariances

Let G be the stationary Gaussian process on \mathbb{R} with covariance function Q , (1.10), and recall that, for any mollifier ϕ , we denote $G_{\phi,\epsilon} = G * \phi_\epsilon$. In this section, we derive the asymptotics of the covariance

$$\mathbb{E}[G_{\phi,\epsilon}(u)G_{\psi,\delta}(v)] = \int_{\mathbb{R}} e^{-2\pi i(u-v)\kappa} \hat{\psi}(\epsilon\kappa) \overline{\hat{\phi}(\delta\kappa)} \hat{Q}(\kappa) d\kappa, \quad (2.34)$$

as $\epsilon, \delta \rightarrow 0$, for all $u, v \in \mathbb{R}$ and for a large class of mollifiers. This is relevant to check that the conditions (2.5), (2.4), and (2.6) are satisfied and apply theorems 2.1 and 2.6 to conclude that the multiplicative chaos measures associated with counting statistics of the CUE and sine process exist and have the same law. For any $\ell, \epsilon > 0$, we define the function

$$Q_\epsilon(x) = -\log\left(\frac{\epsilon}{2\pi} \vee |x|\right) + \log\left(\frac{\epsilon}{2\pi} \vee \sqrt{|\ell^2 - x^2|}\right). \quad (2.35)$$

In the following, we will use the notation $\mathcal{U}_{\mathbf{u}}^{\epsilon \rightarrow 0}$ to specify a function of the variable \mathbf{u} and the parameter $\epsilon > 0$ which is uniformly bounded (by a universal constant) and converge to 0 as $\epsilon \rightarrow 0$ for almost all $\mathbf{u} \in \mathbb{R}^q$.

Recall that we defined $\mathcal{D} = \bigcup_{\alpha > 0} \mathcal{D}_\alpha$, c.f. (1.43), and recall also the definition of the cosine integrals:

$$\text{Cin}(\omega) = \int_0^\omega \frac{1 - \cos(z)}{z} dz \quad \text{and} \quad \text{Ci}(x) = \int_1^\infty \frac{\cos(xt)}{t} dt.$$

The function Cin is entire, while the function Ci is even on \mathbb{R} with the value $+\infty$ at 0, and it turns out that for any $\omega \in \mathbb{R} \setminus \{0\}$,

$$\text{Cin}(\omega) = \text{Ci}(\omega) + \log |\omega| + \gamma,$$

where γ is the Euler constant. In particular, since $\lim_{x \rightarrow \infty} \text{Ci}(x) = 0$, we have for any $\omega \in \mathbb{R}$,

$$\text{Cin}(\omega/\epsilon) = \log^+(\omega/\epsilon) + \gamma + \mathcal{U}_\omega^{\epsilon \rightarrow 0}. \quad (2.36)$$

Lemma 2.11. *Let $\Phi \in L^1(\mathbb{R}_+)$, continuous with $\Phi(0) = 1$, and so that the function $\kappa \mapsto \frac{\Phi(\kappa) - \mathbb{1}_{\kappa \leq 1}}{\kappa}$ is integrable on \mathbb{R}_+ . Then, the function*

$$\mathcal{E}_\Phi(\omega) = \int_0^\infty (1 - \cos(\omega\kappa)) \frac{\Phi(\kappa) - \mathbb{1}_{\kappa \leq 1}}{\kappa} d\kappa$$

is continuously differentiable on \mathbb{R} , and $\lim_{\omega \rightarrow \infty} \mathcal{E}_\Phi(\omega) = \int_0^\infty \frac{\Phi(\kappa) - \mathbb{1}_{\kappa \leq 1}}{\kappa} d\kappa$.

Moreover, for all $\omega \in \mathbb{R}$,

$$\int_0^\infty \frac{1 - \cos(\omega\kappa)}{\kappa} \Phi(\kappa) d\kappa = \text{Cin}(\omega) + \mathcal{E}_\Phi(\omega). \quad (2.37)$$

Proof. All the properties of the function \mathcal{E}_Φ are easy to check. In particular, we have

$$\mathcal{E}'_\Phi(\omega) = \int_0^\infty \sin(\omega\kappa) (\Phi(\kappa) - \mathbb{1}_{\kappa \leq 1}) d\kappa,$$

and the limit of $\mathcal{E}_\Phi(\omega)$ as $\omega \rightarrow \infty$ follows directly from the Riemann-Lebesgue lemma. Finally, the identity (2.37) is an immediate consequence of the definition of the cosine integral. \square

Proposition 2.12. *For any function $\phi, \psi \in \mathcal{D}$, we have for all $u, v \in \mathbb{R}$,*

$$\mathbb{E}[G_{\phi, \epsilon}(u) G_{\psi, \epsilon}(v)] = Q_\epsilon(u - v) + \mathcal{U}_{u, v}^{\epsilon \rightarrow 0}.$$

Proof. Let $\Phi = \Re\{\hat{\phi}\bar{\psi}\}$, $\Psi = \Im\{\hat{\phi}\bar{\psi}\}$, and $\epsilon' = \epsilon/2\pi$. We claim that the function Φ satisfies the assumption of lemma 2.11. Namely, by the Cauchy-Schwarz inequality, $\Phi \in L^1(\mathbb{R})$ and by Plancherel's formula, for any $\kappa \in \mathbb{R}$,

$$\Phi(\kappa) = \iint_{\mathbb{R}^2} \cos(2\pi\kappa x) \phi(x+t) \psi(t) dx dt.$$

In particular, the bound $|e^{i\omega} - 1| \leq 2|\omega|^\alpha$ valid for all $\omega \in \mathbb{R}$ and $0 \leq \alpha \leq 1$ implies that

$$\begin{aligned} \left| \frac{\Phi(\kappa) - 1}{\kappa} \right| &\leq 2\pi\kappa^{\alpha-1} \iint_{\mathbb{R}^2} |x|^\alpha \phi(x+t) \psi(t) dx dt \\ &\leq 2\pi\kappa^{\alpha-1} \int |x|^\alpha \phi(x) dx \int |t|^\alpha \psi(t) dt \end{aligned}$$

Since both $\phi, \psi \in \mathcal{D}_\alpha$ for some $\alpha > 0$, this show that the function $\frac{\Phi(\kappa) - \mathbb{1}_{\kappa \leq 1}}{\kappa}$ is integrable on \mathbb{R}_+ . Similarly, the function $\Psi(\kappa)/\kappa$ is also integral on \mathbb{R}_+ . By the Riemann-Lebesgue lemma, this implies that for any $\omega \in \mathbb{R}$,

$$\int_0^\infty \sin(\epsilon \omega \kappa) \Psi(\kappa) \frac{d\kappa}{\kappa} = \mathcal{U}_{\omega} \underset{\epsilon \rightarrow 0}{\rightarrow} 0. \quad (2.38)$$

By formula (2.34) and by definition of \hat{Q} , we have

$$\mathbb{E}[G_{\phi, \epsilon}(u) G_{\psi, \epsilon}(v)] = 2 \int_0^\infty \Re \left\{ e^{-2\pi i(u-v)\kappa} \hat{\phi}(\epsilon \kappa) \overline{\hat{\psi}(\epsilon \kappa)} \right\} \frac{\sin^2(\pi \ell \kappa)}{\kappa} d\kappa \quad (2.39)$$

$$= \underbrace{\int_0^\infty 2 \cos\left(\frac{(u-v)\kappa}{\epsilon'}\right) \frac{\sin^2(\ell \kappa / 2\epsilon')}{\kappa} \Phi(\kappa) d\kappa}_{:= I_1(u, v, \epsilon')} + \underbrace{\int_0^\infty 2 \sin\left(\frac{(u-v)\kappa}{\epsilon'}\right) \frac{\sin^2(\ell \kappa / 2\epsilon')}{\kappa} \Psi(\kappa) d\kappa}_{:= I_2(u, v, \epsilon')}. \quad (2.40)$$

Using the trigonometric identity

$$-2 \cos(a) \sin^2(b/2) = 1 - \cos a - \frac{1 - \cos(b+a) + 1 - \cos(b-a)}{2},$$

with $a = (u-v)\kappa/\epsilon'$ and $b = \ell\kappa/\epsilon'$, we see that we can apply lemma 2.11. We obtain for all $u, v \in \mathbb{R}$,

$$\begin{aligned} I_1(u, v, \epsilon') &= -\text{Cin}\left(\frac{u-v}{\epsilon'}\right) + \frac{1}{2} \left\{ \text{Cin}\left(\frac{\ell+u-v}{\epsilon'}\right) + \text{Cin}\left(\frac{\ell+v-u}{\epsilon'}\right) \right\} \\ &\quad - \mathcal{E}_\Phi\left(\frac{u-v}{\epsilon'}\right) + \frac{1}{2} \left\{ \mathcal{E}_\Phi\left(\frac{\ell+u-v}{\epsilon'}\right) + \mathcal{E}_\Phi\left(\frac{\ell+v-u}{\epsilon'}\right) \right\}, \end{aligned} \quad (2.41)$$

and the terms in (2.41) combine as the error term $\mathcal{U}_{u,v}$. Moreover, by formula (2.36), this implies that

$$I_1(u, v, \epsilon') = -\log^+\left(\frac{u-v}{\epsilon'}\right) + \frac{1}{2} \left\{ \log^+\left(\frac{\ell+u-v}{\epsilon'}\right) + \log^+\left(\frac{\ell+v-u}{\epsilon'}\right) \right\} + \mathcal{U}_{u,v} \underset{\epsilon \rightarrow 0}{\rightarrow} 0.$$

By definition of the functions \log^+ and Q_ϵ , (2.35), this may be written as

$$I_1(u, v, \epsilon') = Q_\epsilon(u-v) + \mathcal{U}_{u,v} \underset{\epsilon \rightarrow 0}{\rightarrow} 0. \quad (2.42)$$

We can also evaluate the integral $I_2(u, v, \epsilon')$ using a similar argument. Since

$$2 \sin(a) \sin^2(b/2) = \sin a - \sin(a+b)/2 + \sin(a-b)/2,$$

the estimate (2.38) shows that $I_2(u, v, \epsilon') = \mathcal{U}_{u,v}$. Combining this fact with (2.42) and formula (2.40), this completes the proof. \square

Proposition 2.13. *Let $\phi, \psi \in \mathcal{D}$, we have for all $u, v \in \mathbb{R}$,*

$$\mathbb{E}[G_{\phi, \epsilon}(u) G_{\psi, \delta}(v)] = Q_{\epsilon \vee \delta}(u-v) + \mathcal{U}_{u,v} \underset{\delta, \epsilon \rightarrow 0}{\rightarrow} 0,$$

where we can consider the limit as the parameters ϵ and δ converge to 0 in an arbitrary way.

Proof. We let $\epsilon' = \epsilon/2\pi$ and $\delta' = \delta/\epsilon$. Without loss of generality, we suppose that $\delta' \rightarrow \beta \in [0, 1]$ as $\epsilon \rightarrow 0$ (in particular, $\beta = 0$ if we consider successive limits as $\delta \rightarrow 0$ and $\epsilon \rightarrow 0$). By formula (2.39), we have

$$\mathbb{E}[G_{\phi,\epsilon}(u)G_{\psi,\delta}(v)] = 2 \int_0^\infty \Re \left\{ e^{-(u-v)\kappa/\epsilon'} \hat{\phi}(\kappa) \overline{\hat{\psi}(\kappa\delta')} \right\} \frac{\sin^2(\ell\kappa/2\epsilon')}{\kappa} d\kappa.$$

In particular, this implies that

$$|\mathbb{E}[G_{\phi,\epsilon}(u)G_{\psi,\delta}(v)] - \mathbb{E}[G_{\phi,\epsilon}(u)G_{\psi,\epsilon\beta}(v)]| \leq 2 \int_0^\infty \left| \hat{\psi}(\kappa\beta) - \hat{\psi}(\kappa\delta') \right| \left| \hat{\phi}(\kappa) \right| \frac{d\kappa}{\kappa}.$$

Since $\left| \hat{\psi}(\kappa\beta) - \hat{\psi}(\kappa\delta') \right| \leq 4\pi|k|^\alpha |\beta - \delta'|^\alpha \int |x|^\alpha \psi(x) dx$ for any $0 \leq \alpha \leq 1$, there exists a constant $C > 0$ which depends only on $\psi \in \mathcal{D}_\alpha$ so that if $0 < \alpha < 1/2$ is sufficiently small, then

$$|\mathbb{E}[G_{\phi,\epsilon}(u)G_{\psi,\delta}(v)] - \mathbb{E}[G_{\phi,\epsilon}(u)G_{\psi,\epsilon\beta}(v)]| \leq C|\beta - \delta'|^\alpha \int_0^\infty \left| \hat{\phi}(\kappa) \right| \frac{d\kappa}{\kappa^{1-\alpha}}. \quad (2.43)$$

Since $\phi \in L^1 \cap L^2(\mathbb{R})$, by the Cauchy-Schwarz inequality, the RHS of (2.43) is finite and converges to 0 as $\epsilon \rightarrow 0$ (by assumption, $\lim \delta' = \beta$). Thus, it suffices to prove that given $\beta \in [0, 1]$,

$$\mathbb{E}[G_{\phi,\epsilon}(u)G_{\psi,\epsilon\beta}(v)] = Q_\epsilon(u-v) + \mathcal{U}_{u,v}. \quad (2.44)$$

If $\beta > 0$, this follows directly from lemma 2.12, since $G_{\psi,\epsilon\beta} = G_{\psi\beta,\epsilon}$. If $\beta = 0$, since $\hat{\psi}(0) = 1$, we have

$$\mathbb{E}[G_{\phi,\epsilon}(u)G_{\psi,\epsilon\beta}(v)] = \int_0^\infty \Re \left\{ e^{-(u-v)\kappa/\epsilon'} \hat{\phi}(\kappa) \right\} \frac{\sin^2(\ell\kappa/2\epsilon')}{\kappa} d\kappa,$$

and the same computations as in the proof of proposition 2.12 shows that the function $\Re \hat{\phi}$ satisfies the assumptions of lemma 2.11 and the function $\kappa \mapsto \Im \hat{\phi}(\kappa)/\kappa$ is integrable on \mathbb{R}_+ , so that the asymptotics (2.43) hold. \square

Corollary 2.14. *Let $\ell > 0$, G be the stationary Gaussian process on \mathbb{R} with covariance function (1.10), and let $\phi \in \mathcal{D}$. For any $u, v \in \mathbb{R}$, let $T(u, v) = Q(u-v)$, and for any $\epsilon, \delta > 0$, define*

$$T_{\epsilon,\delta}(u, v) = \mathbb{E}[G_{\phi,\epsilon}(u)G_{\phi,\delta}(v)].$$

Then, the function $T_{\epsilon,\delta}$ satisfies the assumption 2.3.

Proof. By formula (2.35), $\lim_{\epsilon \rightarrow 0} Q_\epsilon(x) = Q(x)$ for almost all $x \in \mathbb{R}$, so that the condition (2.5) follows directly from proposition 2.13. Similarly, (2.6) is also an immediate consequence proposition 2.13. To get the upper-bound, we check that if $\epsilon \leq \ell \wedge 1$,

$$Q_\epsilon(x) \leq -\log \left(\frac{\epsilon}{2\pi} \vee |x| \right) + \log \sqrt{\ell^2 + |x|^2}$$

so that we have

$$Q_\epsilon(x) - \log^+(\epsilon^{-1} \wedge |x|^{-1}) \leq \sup \left\{ \log \sqrt{\ell^2 + |x|^2} : |x| \leq 1 \right\} \vee \sup \left\{ -\log |x| + \log \sqrt{\ell^2 + |x|^2} : |x| \geq 1 \right\}.$$

Since the function $x \mapsto \log |x| + \log \sqrt{\ell^2 + |x|^2}$ is decreasing on \mathbb{R}_+ , this shows that

$$Q_{\epsilon \vee \delta}(x) \leq \log^+(|x|^{-1} \wedge \epsilon^{-1} \wedge \delta^{-1}) + \sqrt{\ell^2 + 1}$$

and the condition (2.4) also follows from proposition 2.13. \square

3 Asymptotic analysis

3.1 Proof of proposition 1.11

The goal is to deduce the asymptotics of proposition 1.11 from theorem 1.6 by performing the asymptotics of the solution of the Riemann-Hilbert problem (1.29). Recall that for the sine process, we have

$$f(x) = \begin{pmatrix} e^{i\pi Nx} \\ e^{-i\pi Nx} \end{pmatrix}, \quad g(x) = \begin{pmatrix} e^{i\pi Nx} \\ -e^{-i\pi Nx} \end{pmatrix}, \quad (3.1)$$

so that we look for the asymptotics of the solution of the problem:

- $m(z)$ is analytic on $\mathbb{C} \setminus \mathbb{R}$.
- $m(z)$ satisfies the jump condition:

$$m_+(z) = m_-(z) \begin{pmatrix} 1 + \varphi(x) & -\varphi(x)e^{2\pi i Nx} \\ \varphi(x)e^{-2\pi i Nx} & 1 - \varphi(x) \end{pmatrix}, \quad z \in \mathbb{R} \quad (3.2)$$

- $m(z) \rightarrow I$ as $z \rightarrow \infty$ in $\mathbb{C} \setminus \mathbb{R}$.

Note that we do not emphasize that the matrix m depends on the dimension N to keep the notation simple. Moreover, the solution of (1.29) can be obtained from the solution of (3.2) simply by replacing φ by $\varphi_t = t\varphi$ for all $t \in [0, 1]$. Finally, we will generalize slightly the setting of proposition 1.11 and work with the following assumptions.

Assumption 3.1. *Suppose that φ is a function which is analytic in the strip $|\Im z| < \delta$, so that $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ and $\varphi \in L^1 \cap L^\infty(\mathbb{R} \pm is)$ for all $|s| \leq \delta/2$. In particular, we define*

$$C_\varphi = \sup \{ \|\varphi\|_{L^\infty(\mathbb{R} \pm is)} \vee \|\varphi\|_{L^1(\mathbb{R} \pm is)} : s \leq \delta/2 \}.$$

Let us also assume that there exists a constant $c > 0$ so that

$$\sup \{ |\varphi(z) - x| : |\Im z| < \delta, x < -1 \} \geq c. \quad (3.3)$$

In particular, the function $\psi = \log(1 + \varphi)$ is also analytic in the strip $|\Im z| < \delta$. Finally, we assume that $\psi \in L^1 \cap L^\infty(\mathbb{R})$ and we let $C_\psi = c^{-1} \exp \|\psi\|_{L^\infty(\mathbb{R})}$.

Lemma 3.1. *Suppose that the function φ satisfies the assumption 3.1 and that $C_\varphi C_\psi^2 e^{-\pi \delta N} \rightarrow 0$ as $N \rightarrow \infty$. Then, the solution of the Riemann-Hilbert problem (3.2) satisfies for all $x \in \mathbb{R}$,*

$$m_+(x) = R(x) \begin{pmatrix} e^{\mathcal{C}(\psi)_+(x)} & -\frac{\varphi(x)}{1+\varphi(x)} e^{\mathcal{C}(\psi)_+(x)+2\pi i Nx} \\ 0 & e^{-\mathcal{C}(\psi)_+(x)} \end{pmatrix}, \quad (3.4)$$

where $\mathcal{C}(\psi)$ denotes the Cauchy transform of the function $\psi = \log(1 + \varphi)$ and the 2×2 matrix $R(z)$ is analytic in the strip $|\Im z| < \delta/4$ and satisfies the bound:

$$\|R(z) - I\| \ll \delta^{-1/2} C_\varphi C_\psi^2 e^{-\pi \delta N}. \quad (3.5)$$

The proof of lemma 3.1 will be given at the end of this section and it follows closely the proof of the Strong Szegő theorem given by Deift in [16, Example 3]. Given this result, let us first complete the proof of proposition 1.11.

Proof of proposition 1.11. First of all, we claim that, if the function $h(z)$ satisfies the assumption 1.2, then the function $\varphi_t(z) = t(e^{h(z)} - 1)$ satisfies the assumption 3.1 for all $t \in [0, 1]$. Indeed, we have

$$|\varphi_t(z)| \leq \left| \int_0^{h(z)} e^w dw \right| \leq C_\infty C_1$$

and for all $|s| \leq \delta/2$,

$$\int_{\mathbb{R}} |\varphi_t(x + is)| dx \leq C_{\infty} \int_{\mathbb{R}} |h(x + is)| dx \leq C_{\infty} C_1. \quad (3.6)$$

On the other hand, the condition $|\Im h(z)| < \alpha$ guarantees we can choose $c = |\sin \alpha|/2$ in (3.3). Thus, for all $t \in [0, 1]$, the functions $\psi_t(z) = \log(1 - t + te^{h(z)})$ are analytic in the strip $|\Im z| < \delta$, and since the log function is increasing on \mathbb{R}_+ , we have for all $x \in \mathbb{R}$,

$$\psi_t(x) \vee 0 \leq \log(e^{h(x)} \vee 1) = h(x) \vee 0 \quad \text{and} \quad -\psi_t(x) \wedge 0 \leq -\log(e^{h(x)} \wedge 1) \leq h(x) \wedge 0.$$

These inequalities show that

$$|\psi_t(x)| \leq |h(x)| \quad (3.7)$$

and it follows that $\psi_t \in L^1 \cap L^{\infty}(\mathbb{R})$ (in fact there is equality in (3.7) if and only if $t = 1$, in which case $\psi_1 = h$). The bottom line is that, for any $t \in [0, 1]$, the function $\varphi_t(z) = t(e^{h(z)} - 1)$ satisfies assumption 3.1 with

$$C_{\varphi_t} = C_{\infty} C_1 \quad \text{and} \quad C_{\psi_t} = \frac{2C_{\infty}}{|\sin \alpha|}. \quad (3.8)$$

In the rest of the proof, we will denote for all $t \in [0, 1]$ and for all $x \in \mathbb{R}$,

$$\psi_t(z) = \log(1 + t\varphi) = \log(1 - t + te^{h(z)}) \quad \text{and} \quad H_t(x) = \mathcal{C}(\psi_t)_+(x).$$

By theorem 1.6, formula (3.4) implies that

$$F_t(x) = m_+(x)f(x) = R_t(x)\tilde{F}_t(x) \quad \text{where} \quad \tilde{F}_t(x) = \begin{pmatrix} \frac{e^{i\pi Nx + H_t(x)}}{1 + t\varphi(x)} \\ e^{-i\pi Nx - H_t(x)} \end{pmatrix},$$

and

$$G_t(x) = (m_+^{-1})^*(x)g(x) = (R_t(x)^{-1})^*\tilde{G}_t(x) \quad \text{where} \quad \tilde{G}_t(x) = \begin{pmatrix} e^{i\pi Nx + H_t(x)} \\ -\frac{e^{-i\pi Nx - H_t(x)}}{1 + t\varphi(x)} \end{pmatrix}.$$

Since the function $h' \in L^1(\mathbb{R})$, the function H_t is continuously differentiable on \mathbb{R} for all $t \in [0, 1]$. Moreover, by the Plemelj-Sokhotski formula (A.6), we obtain

$$2H'_t = \frac{t\varphi'}{1 + t\varphi} + i\mathcal{H}(\psi'_t). \quad (3.9)$$

So, if we differentiate the expression of $\tilde{F}_t(x)$, we obtain for all $x \in \mathbb{R}$,

$$\tilde{F}'_t(x) = \begin{pmatrix} (\pi iN + H'_t(x) - \frac{t\varphi'(x)}{1 + t\varphi(x)}) \frac{1}{1 + t\varphi(x)} e^{\pi iNx + H_t(x)} \\ -(\pi iN + H'_t(x)) e^{-\pi iNx - H_t(x)} \end{pmatrix}.$$

Since, $F_t^* G_t = 0$, this shows that

$$\begin{aligned} \left(\frac{d\sqrt{\varphi}F_t}{dx}(x) \right)^* \left(\sqrt{\varphi(x)}G_t(x) \right) &= \varphi(x)F'_t(x)^*G_t(x) \\ &= \frac{\varphi(x)}{1 + t\varphi(x)} \left(-2\pi iN + 2\overline{H'_t(x)} - \frac{t\varphi'(x)}{1 + t\varphi(x)} \right) + \varphi(x) \left(U_t(x)\tilde{F}_t(x) \right)^* \tilde{G}_t(x), \end{aligned} \quad (3.10)$$

where $U_t(x) = R_t(x)^{-1}R'_t(x)$. Since the matrix $R_t(z)$ is analytic in the strip $|\Im z| < \delta/4$, by Cauchy's formula, $\|R'_t(x)\| \leq \delta^{-1}\|R_t(x) - I\|$. Thus, the estimate (3.5) and (3.8) show that

$$\|U_t(x)\| \ll \delta^{-3/2}C_{\varphi_t}C_{\psi_t}e^{-\pi\delta N} \ll \frac{C_{\infty}^3C_1}{\delta^{3/2}|\sin \alpha|}e^{-\pi\delta N}.$$

On the other hand, $\Re\{H_t\} = \frac{1}{2} \log(1 + t\varphi) = \frac{\psi_t}{2}$ so that, by the estimate (3.7), for all $x \in \mathbb{R}$ and all $t \in [0, 1]$,

$$\|\tilde{F}_t(x)\| \leq \sqrt{C_\infty} \quad \text{and} \quad \|\tilde{G}_t(x)\| \leq C_\infty^{3/2}.$$

This proves that the last term in formula (3.10) is bounded by

$$\left| \varphi(x) \left(U_t(x) \tilde{F}_t(x) \right)^* \tilde{G}_t(x) \right| \ll |\varphi(x)| \frac{C_\infty^5 C_1}{\delta^{3/2} |\sin \alpha|} e^{-\pi \delta N}. \quad (3.11)$$

By (3.6), the RHS of (3.11) is integrable on $\mathbb{R} \times [0, 1]$ and its L^1 -norm contributes to the error term in formula (1.46). Therefore, by formula (1.28) and (3.10), in order to complete the proof, it remains to show that

$$\int_0^1 \int_{\mathbb{R}} \frac{\varphi(x)}{1 + t\varphi(x)} \left(2\pi i N - 2\overline{H'_t(x)} + \frac{t\varphi'(x)}{1 + t\varphi(x)} \right) dx dt = 2\pi i N \int h(x) dx + i\pi \|h\|_{H^{1/2}(\mathbb{R})}^2. \quad (3.12)$$

This is a rather easy computation that we split in two steps. First observe that, since $\varphi(x) = e^{h(x)} - 1$, we have

$$\int_0^1 \frac{\varphi(x)}{1 + t\varphi(x)} dt = [\log(1 + t\varphi(x))]_{t=0}^1 = h(x).$$

By Fubini's theorem, this yields the first term in formula (3.12). Secondly, by formula (3.9), we have

$$\frac{\varphi}{1 + t\varphi} \left(2H'_t - \frac{t\varphi'}{1 + t\varphi} \right) = i \frac{d\psi_t}{dt} \mathcal{H} \left(\frac{d\psi_t}{dx} \right).$$

By Plancherel's formula and the linearity of the Fourier transform, this implies that

$$\begin{aligned} \int_{\mathbb{R}} \frac{d\psi_t}{dt}(x) \mathcal{H} \left(\frac{d\psi_t}{dx} \right)(x) dx &= 2\pi \int_{\mathbb{R}} |\kappa| \widehat{\frac{d\psi_t}{dt}}(\kappa) \overline{\widehat{\psi_t}(\kappa)} d\kappa \\ &= \pi \frac{d}{dt} \left(\int_{\mathbb{R}} |\kappa| \widehat{\psi_t}(\kappa) \overline{\widehat{\psi_t}(\kappa)} d\kappa \right) \\ &= \|\psi_t\|_{H^{1/2}(\mathbb{R})}^2 \end{aligned}$$

Note that we can pull the differential $\frac{d}{dt}$ out of the integral because both functions $\psi_t, \frac{d\psi_t}{dt} \in L^2(\mathbb{R})$ by our assumptions. We conclude that

$$\int_0^1 \int_{\mathbb{R}} \frac{\varphi(x)}{1 + t\varphi(x)} \left(2H'_t(x) - \frac{t\varphi'(x)}{1 + t\varphi(x)} \right) dx dt = i\pi \left(\|\psi_1\|_{H^{1/2}(\mathbb{R})}^2 - \|\psi_0\|_{H^{1/2}(\mathbb{R})}^2 \right).$$

Since $\psi_1 = h$ and $\psi_0 = 0$, this yields the second term in formula (3.12). \square

In order to get the uniform estimate (3.5) in lemma 3.1, we will need the following correspondence between a well-posed RHP and a singular integral equation.

Theorem 3.2 (Kuijlaars [42], Theorem 3.1). *Let Σ be an oriented contour in the complex plane and let $\Delta \in L^2 \cap L^\infty(\Sigma)$ be a $n \times n$ matrix. Suppose that R is a $n \times n$ matrix which is analytic in $\mathbb{C} \setminus \Sigma$ and satisfies*

$$\begin{cases} R_+ = R_-(I + \Delta) & \text{on } \Sigma \\ R(z) \rightarrow I & \text{as } z \rightarrow \infty \end{cases}. \quad (3.13)$$

We associate to Δ an operator \mathcal{C}_Δ acting on $n \times n$ matrices in $L^2(\Sigma)$ and defined by

$$\mathcal{C}_\Delta Y(z) = \lim_{w \rightarrow z_+} \frac{1}{2\pi i} \int_{\Sigma} \frac{Y(s) \Delta(s)}{s - w} ds.$$

Suppose that $\|\Delta\|_{L^\infty(\Sigma)}$ is sufficiently small so that the equation

$$X - C_\Delta X = C_\Delta I \quad (3.14)$$

has a unique solution $X \in L^2(\Sigma)$. Then, there exists a constant $C > 0$ which only depends on the contour Σ such that

$$\|X\|_{L^2(\Sigma)} \leq \frac{C\|\Delta\|_{L^2(\Sigma)}}{1 - C\|\Delta\|_{L^\infty(\Sigma)}}, \quad (3.15)$$

and the RHP (3.13) has a unique solution which is given by

$$R(z) = I + \frac{1}{2\pi i} \int_\Sigma \frac{(I + X(s))\Delta(s)}{s - z} ds$$

for any $z \in \mathbb{C} \setminus \Sigma$.

Armed with the above theorem, we can construct the solution of the Riemann-Hilbert problem (3.2) and check the estimate (3.5).

Proof of lemma 3.1. If $\psi(x) = \log(1 + \varphi(x))$, then we can rewrite the jump matrix as

$$v(x) = \begin{pmatrix} e^{\psi(x)} & -\varphi(x)e^{2\pi i N x} \\ \varphi(x)e^{-2\pi i N x} & (1 - \varphi(x)^2)e^{-\psi(x)} \end{pmatrix}.$$

In order to apply the Deift-Zhou steepest method, we use the decomposition:

$$v(x) = \underbrace{\begin{pmatrix} 1 & 0 \\ \varphi(x)e^{-\psi(x)-2\pi i N x} & 1 \end{pmatrix}}_{= \tilde{A}(x)} \underbrace{\begin{pmatrix} e^{\psi(x)} & 0 \\ 0 & e^{-\psi(x)} \end{pmatrix}}_{= \Psi(x)} \underbrace{\begin{pmatrix} 1 & -\varphi(x)e^{-\psi(x)+2\pi i N x} \\ 0 & 1 \end{pmatrix}}_{= A(x)}. \quad (3.16)$$

By assumption, the matrices A and \tilde{A} are analytic in the domain $|\Im z| < \delta$ and we can define a matrix M on $\mathbb{C} \setminus \{\mathbb{R} \cup \Gamma_\pm\}$ by setting:

$$\begin{array}{c} M(z) = m(z) \\ \hline M(z) = m(z)A^{-1}(z) \\ \hline M(z) = m(z)\tilde{A}(z) \\ \hline \end{array} \begin{array}{l} \Gamma_+ = \mathbb{R} + i\delta/2 \\ \mathbb{R} \\ \Gamma_- = \mathbb{R} - i\delta/2 \end{array}.$$

We deduce from the Riemann-Hilbert problem (3.2), that the matrix M has the following properties:

- $M(z)$ is analytic on $\mathbb{C} \setminus (\Gamma_\pm \cup \mathbb{R})$.
- $M(z)$ satisfies the following jump conditions

$$\begin{cases} M_+(z) = M_-(z)A(z), & z \in \Gamma_+ \\ M_+(x) = M_-(x)\Psi(x), & x \in \mathbb{R} \\ M_+(z) = M_-(z)\tilde{A}(z), & z \in \Gamma_- \end{cases} \quad (3.17)$$

- $M(z) \rightarrow I$ as $z \rightarrow \infty$.

Since $\psi \in L^1(\mathbb{R})$, its Cauchy transform is well-defined, c.f. (A.5), and we claim that the global parametrix is given by

$$P(z) = \begin{pmatrix} e^{\mathcal{C}(\psi)(z)} & 0 \\ 0 & e^{-\mathcal{C}(\psi)(z)} \end{pmatrix}. \quad (3.18)$$

Indeed, it is straightforward to check using formula (A.6), that it solves the RHP:

- $P(z)$ is analytic on $\mathbb{C} \setminus \mathbb{R}$.
- $P(z)$ satisfies the condition $P_+ = P_- \Psi$ on \mathbb{R}
- $P(z) \rightarrow \mathbf{I}$ as $z \rightarrow \infty$.

The matrix $P(z)$ is invertible on $\mathbb{C} \setminus \mathbb{R}$ and, if we let $R = MP^{-1}$, then the matrix R solves the RHP:

- $R(z)$ is analytic on $\mathbb{C} \setminus (\mathbb{R} \cup \Gamma_{\pm})$.
- If we let $\Delta = PAP^{-1} - \mathbf{I}$ and $\tilde{\Delta} = P\tilde{A}P^{-1} - \mathbf{I}$, then $R(z)$ satisfies the jump conditions

$$\begin{cases} R_+ = R_-(\mathbf{I} + \Delta) & \text{on } \Gamma_+ \\ R_+ = R_-(\mathbf{I} + \tilde{\Delta}) & \text{on } \Gamma_- \end{cases} \quad (3.19)$$

- $R(z) \rightarrow \mathbf{I}$ as $z \rightarrow \infty$.

Moreover observe that by formulae (3.16) and (3.18),

$$\begin{aligned} \Delta(z) &= \begin{pmatrix} 0 & -\varphi(z)e^{-\psi(z)+2\pi iNz+2\mathcal{C}(\psi)(z)} \\ 0 & 0 \end{pmatrix}, \quad \forall z \in \Gamma_+ \\ \tilde{\Delta}(z) &= \begin{pmatrix} 0 & 0 \\ \varphi(z)e^{-\psi(z)-2\pi iNz-2\mathcal{C}(\psi)(z)} & 0 \end{pmatrix}, \quad \forall z \in \Gamma_-. \end{aligned}$$

The function ψ is real valued on \mathbb{R} and for any $z \in \Gamma_{\pm}$, we have

$$\Re\{\mathcal{C}(\psi)(z)\} = \frac{\delta}{4\pi} \int_{\mathbb{R}} \frac{\psi(x)}{(\Re z - x)^2 + \delta^2/4} dx,$$

so that

$$|\Re\{\mathcal{C}(\psi)(z)\}| \leq \|\psi\|_{L^\infty(\mathbb{R})}/2.$$

Combined with the assumption 3.1, this implies that for any $z \in \Gamma_+$,

$$\|\Delta(z)\| = \left| \frac{\varphi(z)}{1 + \varphi(z)} \right| e^{\Re\{2\pi iNz+2\mathcal{C}(\psi)(z)\}} \leq |\varphi(z)| C_\psi^2 e^{-\pi\delta N},$$

so that the matrix $\Delta \in L^1 \cap L^\infty(\Gamma_+)$ and

$$\|\Delta\|_{L^\infty(\Gamma_+)} \vee \|\Delta\|_{L^1(\Gamma_+)} \leq C_1 C_\psi^2 e^{-\pi\delta N}. \quad (3.20)$$

Note that, since $\tilde{\Delta}(z) = -\Delta(\bar{z})^*$ for any $z \in \Gamma_-$, the matrix $\tilde{\Delta}$ also satisfies the estimate (3.20). Hence, by theorem 3.2, we obtain that the solution of the RHP (3.19) is unique and is given by

$$R(z) = \mathbf{I} + \frac{1}{2\pi i} \int_{\Gamma_+} \frac{(\mathbf{I} + X(s))\Delta(s)}{s - z} ds + \frac{1}{2\pi i} \int_{\Gamma_-} \frac{(\mathbf{I} + \tilde{X}(s))\tilde{\Delta}(s)}{s - z} ds \quad (3.21)$$

where the 2×2 matrices X and \tilde{X} solve appropriate singular integral equations, c.f. (3.14). Moreover, a simple estimate using the Cauchy-Schwarz inequality shows that for any $z \in \mathbb{C}$ such that $|\Im z| \leq \delta/4$,

$$\left\| \frac{1}{2\pi i} \int_{\Gamma_+} \frac{(\mathbf{I} + X(s))\Delta(s)}{s - z} ds \right\| \leq \frac{1}{2\sqrt{\pi\delta}} (\|\Delta\|_{L^2(\Gamma_+)} + \|\Delta\|_{L^\infty(\Gamma_+)} \|X\|_{L^2(\Gamma_+)}).$$

Hence, as $\|\Delta\|_{L^\infty(\Gamma_+)} \rightarrow 0$ as $N \rightarrow \infty$, combining the bounds (3.15) and (3.20), we have proved that

$$\left\| \frac{1}{2\pi i} \int_{\Gamma_+} \frac{(I + X(s))\Delta(s)}{s - z} ds \right\| \ll \delta^{-1/2} C_\psi^2 C_\varphi e^{-\pi\delta N}.$$

A similar estimate holds for the integral over Γ_- and, by formula (3.21), this proves the bound (3.5). Going back to the original problem, we see that for all $x \in \mathbb{R}$, $m_+(x) = M_+(x)A(x)$ where $M_+(x) = R(x)P_+(x)$. Finally, we deduce (3.4) from formulae (3.16) and (3.18). \square

3.2 Borodin-Okounkov formula

The goal of this section is to prove the strong Gaussian approximation for the CUE. We work with the statistic (1.5) and it will be convenient to use the notation

$$h_{\mathbf{u}, \epsilon}(x) = \gamma\pi \sum_{k=1}^q t_k \int_{u_k - l/2}^{u_k + l/2} \phi_{\epsilon_k}(x - t) dt \quad (3.22)$$

We shall suppose that the mollifier ϕ belongs to the Schwarz class, *i.e.* ϕ is a smooth function such that for all $\eta > 0$ we have $\phi(x) = O(|x|^{-\eta})$ as $x \rightarrow \pm\infty$. The main goal of this section will be to prove

Proposition 3.3. *Consider a CUE matrix U of size $N \times N$ and mesoscopic scale $0 < \alpha < 1$. Take $\epsilon^* = \min\{\epsilon_1, \dots, \epsilon_q\}$. Let $\delta > 0$ and suppose that we have the following bound $N^\alpha/(N\epsilon^*) = O(N^{-\delta})$. Then for all Schwarz ϕ , the following Gaussian estimate holds as $N \rightarrow \infty$*

$$\mathbb{E} \left[\exp \left(\sum_{j=1}^N h_{\mathbf{u}, \epsilon}(N^\alpha \theta_j) - \mathbb{E} \left(\sum_{j=1}^N h_{\mathbf{u}, \epsilon}(N^\alpha \theta_j) \right) \right) \right] = \exp \left(\frac{\|h_{\mathbf{u}, \epsilon}\|_{H^{1/2}}^2}{2} \right) (1 + \mathcal{E}_N^S) e^{\mathcal{E}_N^{\text{glob}}} \quad (3.23)$$

where for any $\eta > 0$, the smoothing error term satisfies the bound

$$|\mathcal{E}_N^S| \leq C\gamma^2 \|t\|_1^{\eta_2} N^{-\eta} \exp(\gamma^2 \|t\|_1^{\eta_2} N^{-\eta}) \quad (3.24)$$

for some $\eta_2 > 0$, while the global error term satisfies

$$|\mathcal{E}_N^{\text{glob}}| \leq C\|t\|_1^2 \log(L(N)/\epsilon^*) L(N) N^{-\alpha} \quad (3.25)$$

uniformly in compact subsets of the parameters $u_1, \dots, u_q \in \mathbb{R}$.

Remark 3.4. *This shows that the CUE satisfies the strong Gaussian approximation, *i.e.* the assumption (2.7) is satisfied, which is enough to complete the proofs in Sections 2.1 and 2.2 for the CUE.*

Remark 3.5. *The error (3.24) is called the smoothing error because it fails when $\epsilon \rightarrow 0$ quickly enough that $\epsilon \sim N^\alpha/N$. Beyond this regime the asymptotics in (3.23) are no longer valid and one enters the transition regime of Fisher-Hartwig asymptotics, see [41, 17] for a review. On the other hand, (3.25) is controlled by the relative numbers of eigenvalues sampled by the statistic and is only small in the mesoscopic regime $0 < \alpha < 1$. Furthermore, when the size of the interval $L(N) \rightarrow \infty$ we need this to happen slowly enough that $L(N) = N^\beta$ with $0 < \beta < \alpha$. Then the global error term is $o(1)$ and we obtain theorem 1.3 from the covariance computation of Section 2.3.*

As for the sine process and proposition 1.11, the proof of proposition 3.3 is also based on the existence of integrable and determinantal structures in the CUE. In fact the Laplace transform of (1.5) could also be written as a Fredholm determinant which could then be analysed with an appropriate Riemann-Hilbert problem (as was done in [16] for the global scale $\alpha = 0$ and $\epsilon > 0$)

fixed). However, for the CUE we will give a more elementary proof that does not involve Riemann-Hilbert techniques, but instead relies on an ‘algebraic miracle’ known as the Borodin-Okounkov formula. Unlike our Riemann-Hilbert computation, the proof we give here does not impose any analyticity condition on the mollifier ϕ , but we do require it to be smooth in general. Another crucial difference is that the CUE has a macroscopic regime (defined by (1.5) with $\alpha = 0$) which has no analogue for the sine process.

When working with the CUE it is convenient to expand such functions in a Fourier series. Therefore in what follows we are going to work with the periodisation

$$h_{\mathbf{u},\epsilon}^{(2\pi)}(\theta) = \sum_{a=-\infty}^{\infty} h_{\mathbf{u},\epsilon}(N^\alpha(\theta + 2\pi a)). \quad (3.26)$$

Then $h_{\mathbf{u},\epsilon}^{(2\pi)}(\theta + 2\pi) = h_{\mathbf{u},\epsilon}^{(2\pi)}(\theta)$ is 2π -periodic. The periodisation has the convenient property that its Fourier coefficients are given explicitly in terms of the Fourier transform of $h_{\mathbf{u},\epsilon}$.

$$\frac{1}{2\pi} \int_0^{2\pi} h_{\mathbf{u},\epsilon}^{(2\pi)}(\theta) e^{-ik\theta} d\theta = N^{-\alpha} \int_{\mathbb{R}} h_{\mathbf{u},\epsilon}(2\pi x) e^{-2i\pi kx N^{-\alpha}} dx \quad (3.27)$$

Furthermore, the linear statistics of $h_{\mathbf{u},\epsilon}^{(2\pi)}(\theta)$ are uniformly close to those of $h_{\mathbf{u},\epsilon}(\theta)$. This follows from the rapid decay of ϕ , since the difference between the two is deterministically bounded by

$$\sum_{|a|>0} \sum_{j=1}^N \sum_{k=1}^q \epsilon_k^{\eta-1} \int_{x_j - u_k - l/2}^{x_j - u_k + l/2} |s|^{-\eta} ds \leq \|t\| cN (\epsilon_*/N^\alpha)^\eta \quad (3.28)$$

where $x_j = N^\alpha(\theta_j + 2\pi a)$ and we used that $lN^{-\alpha} \rightarrow 0$ as $N \rightarrow \infty$. Choosing $\eta > 0$ large enough we can always ensure that (3.28) goes to 0 as $N \rightarrow \infty$. Hence it will suffice to always work with the periodisation.

Proposition 3.6 (Macroscopic approximation). *Define the quantity*

$$E_{\mathbf{u},\epsilon} := \frac{1}{(2\pi)^2} N^{-\alpha} \sum_{k=1}^{\infty} k N^{-\alpha} |\hat{h}_{\mathbf{u},\epsilon}(k/(2\pi N^\alpha))|^2 \quad (3.29)$$

and suppose that the hypotheses of proposition 3.3 are satisfied. Then we have the Gaussian estimate

$$\mathbb{E} \left[\exp \left(\sum_{j=1}^N h_{\mathbf{u},\epsilon}^{(2\pi)}(\theta_j) - \mathbb{E} \left(\sum_{j=1}^N h_{\mathbf{u},\epsilon}^{(2\pi)}(\theta_j) \right) \right) \right] = \exp(E_{\mathbf{u},\epsilon}) (1 + \mathcal{E}_N^S) \quad (3.30)$$

where the error term \mathcal{E}_N^S satisfies (3.24) and is uniform in the variables u_1, \dots, u_q varying in compact subsets of \mathbb{R} .

To prove proposition 3.6, the main idea is to exploit the fact that for the CUE, the left-hand side of (3.30) can be written exactly as an $N \times N$ Toeplitz determinant involving the Fourier coefficients of the periodic function $w_{\mathbf{u},\epsilon}(\theta) = e^{h_{\mathbf{u},\epsilon}^{(2\pi)}(\theta)}$. The representation as a determinant follows from the following well known but remarkable chain of equalities for the (un-centered) left-hand

side of (3.23):

$$\mathbb{E} \left[\exp \left(\sum_{j=1}^q X_{N, \epsilon_N}(u_j) \right) \right] \quad (3.31)$$

$$= \frac{1}{(2\pi)^N N!} \int_0^{2\pi} \dots \int_0^{2\pi} \prod_{j=1}^N w(\theta_j) \prod_{1 \leq p < q \leq N} |e^{i\theta_p} - e^{i\theta_q}|^2 d\theta_1 \dots d\theta_q \quad (3.32)$$

$$= \frac{1}{(2\pi)^N N!} \int_0^{2\pi} \dots \int_0^{2\pi} \det\{w(\theta_j) e^{i(k-1)\theta_j}\}_{j,k=1}^N \det\{e^{-i(k-1)\theta_j}\}_{j,k=1}^N d\theta_1 \dots d\theta_q \quad (3.33)$$

$$= \det \left\{ \frac{1}{2\pi} \int_0^{2\pi} w(\theta) e^{-i(k-j-2)\theta} d\theta \right\} \quad (3.34)$$

$$= \det T_N(w) \quad (3.35)$$

where $T_N(w)$ is the $N \times N$ Toeplitz matrix $\{\hat{w}_{k-j-2}\}_{j,k=1}^N$. That (3.31) equals (3.32) is a consequence of the Weyl integration formula for the unitary group. Then (3.33) writes the product of differences in (3.32) as a product of two determinants and finally (3.34) is a consequence of the Andrejef identity.

Thus our task will be to calculate the asymptotics of the Toeplitz determinant in (3.35) as $N \rightarrow \infty$. The function $w(\theta)$ is called the *symbol*. For smooth and N -independent symbols, such asymptotics are well known from the strong Szegő limit theorem. However, our symbol is N -dependent and furthermore the quantity $E_{\mathbf{u}, \epsilon}$ is divergent in the limit $N \rightarrow \infty$. This is due to the fact that our symbol becomes discontinuous as $\epsilon_N \rightarrow 0$ and therefore does not belong to the $H^{1/2}$ -space. The following formula is our main tool for establishing the strong Gaussian approximation in the CUE.

Theorem 3.7 (Borodin-Okounkov formula). *Let $w(\theta)$ a periodic function on the interval $\theta \in [0, 2\pi)$ and that $\log w(\theta)$ has Fourier coefficients \hat{L}_k . Suppose that we have the expansion*

$$\log w(\theta) = \sum_{k=-\infty}^{\infty} \hat{L}_k e^{ik\theta}, \quad \sum_{k=1}^{\infty} k |\hat{L}_k|^2 < \infty \quad (3.36)$$

In terms of the quantities $b(\theta) = 1/c(\theta)$ and

$$c(\theta) = \exp \left(\sum_{k=1}^{\infty} \hat{L}_k e^{ik\theta} - \sum_{k=1}^{\infty} \hat{L}_{-k} e^{-ik\theta} \right) \quad (3.37)$$

we have the following identity

$$\frac{\det(T_N(w))}{\exp \left(N \hat{L}_0 + \sum_{k=1}^{\infty} k |\hat{L}_k|^2 \right)} = \det(I - R_N H(b) H(\tilde{c}) R_N) \quad (3.38)$$

where R_N is the projection operator on $\ell^2(N+1, N+2, \dots)$ and $H(b)$, $H(\tilde{c})$ are infinite Hankel matrices corresponding to the sequence of Fourier coefficients of $b(\theta)$ and $c(\theta)$,

$$H(b) = \begin{pmatrix} b_1 & b_2 & b_3 \dots \\ b_2 & b_3 & b_4 \dots \\ b_3 & b_4 & b_5 \dots \\ \dots & \dots & \dots \end{pmatrix}, \quad H(\tilde{c}) = \begin{pmatrix} c_{-1} & c_{-2} & c_{-3} \dots \\ c_{-2} & c_{-3} & c_{-4} \dots \\ c_{-3} & c_{-4} & c_{-5} \dots \\ \dots & \dots & \dots \end{pmatrix} \quad (3.39)$$

Proof. There are many proofs in the literature. For a comprehensive proof under the conditions mentioned here see [64], theorem 6.2.14, which also has a detailed historical account of formula (3.38). \square

Note that here $\log w(\theta) = h_{\mathbf{u}, \epsilon}^{(2\pi)}$ and the quantity $\sum_{k=1}^{\infty} k \hat{L}_k$ in (3.38) is precisely $E_{\mathbf{u}, \epsilon}$ in (3.29). Furthermore, an easy computation shows that $N \hat{L}_0 = \mathbb{E}(\sum_{j=1}^N h_{\mathbf{u}, \epsilon}^{(2\pi)}(\theta_j))$ which corresponds to a re-centering by the expectation. Hence the Borodin-Okounkov formula (3.38) implies

$$\mathbb{E} \left[\exp \left(\sum_{j=1}^N h_{\mathbf{u}, \epsilon}^{(2\pi)}(\theta_j) - \mathbb{E} \left(\sum_{j=1}^N h_{\mathbf{u}, \epsilon}^{(2\pi)}(\theta_j) \right) \right) \right] = \exp(E_{\mathbf{u}, \epsilon}) \det(I - R_N H(b) H(\tilde{c}) R_N). \quad (3.40)$$

The next lemma gives us the necessary estimate on the above determinant, thus concluding the proof of proposition 3.6.

Lemma 3.8. *Suppose the hypotheses of proposition 3.3 are satisfied. Then for any $\eta > 0$, there exists $\eta_2 > 0$ such that we have the following estimate*

$$|\det(1 - R_N H(b) H(\tilde{c}) R_N) - 1| \leq C \gamma^2 \|t\|_1^{\eta_2} N^{-\eta} \exp(\gamma^2 \|t\|_1^{\eta_2} N^{-\eta}) \quad (3.41)$$

uniformly for variables u_1, \dots, u_q belonging to a compact subset of \mathbb{R} .

Proof. The following inequality is an easy consequence of standard properties of the Fredholm determinant

$$|\det(I + A) - 1| \leq e^{\|A\|_1} - 1 \quad (3.42)$$

In our case $A = -R_N H(b) H(\tilde{c}) R_N = -|H(\tilde{c}) R_N|^2 \leq 0$ using the fact that $H(b) = H(\tilde{c})^\dagger$, so $-A$ is a positive operator. Thus we can write the bound (3.42) in terms of the Hilbert-Schmidt norm

$$|\det(I - R_N H(b) H(\tilde{c}) R_N) - 1| \leq e^{\|H(\tilde{c}) R_N\|_2^2} - 1 \leq \|H(\tilde{c}) R_N\|_2^2 e^{\|H(\tilde{c}) R_N\|_2^2} \quad (3.43)$$

which can be computed explicitly (see 6.2.57 in [64])

$$\|H(\tilde{c}) R_N\|_2^2 = \sum_{k=1}^{\infty} k |c_{k+N}|^2. \quad (3.44)$$

We now proceed to estimate the Fourier coefficients of the function $c(\theta)$ in (3.37), which clearly satisfies $|c(\theta)| = 1$. We have

$$c_{k+N} = \frac{1}{2\pi} \int_0^{2\pi} e^{-i(k+N)\theta} e^{2iu(\theta)} d\theta \quad (3.45)$$

where

$$u(\theta) = \Im \left\{ \sum_{k=1}^{\infty} \hat{L}_k e^{ik\theta} - \sum_{k=1}^{\infty} \hat{L}_{-k} e^{-ik\theta} \right\}. \quad (3.46)$$

The idea is to exploit cancellations in the integral (3.45) for large N coming from rapid oscillations of the factor $e^{-i(k+N)\theta}$. To this end, we integrate by parts p times, obtaining

$$c_{k+N} = \frac{2^p}{2\pi} \frac{1}{(k+N)^p} \int_0^{2\pi} e^{-i(k+N)\theta} e^{-u(\theta)} \left(\frac{d^p}{d\theta^p} e^{u(\theta)} \right) e^{2iu(\theta)} d\theta \quad (3.47)$$

The function $e^{-u(\theta)} \left(\frac{d^p}{d\theta^p} e^{u(\theta)} \right)$ is a polynomial in $u(\theta)$ and all its derivatives up to order p . We have the explicit formula

$$e^{-u(\theta)} \left(\frac{d^p}{d\theta^p} e^{u(\theta)} \right) = \sum_{m=1}^p \sum_{\substack{k_1+2k_2+\dots+pk_p=p \\ k_1+k_2+\dots+k_p=m}} \sum_{\vec{k}} c_{p, \vec{k}} \prod_{l=1}^p u^{(l)}(\theta)^{k_l} \quad (3.48)$$

where $c_{p,\vec{k}}$ are some combinatorial coefficients. We now proceed to estimate $u^{(l)}(\theta)$. Clearly the interchange of derivative and sum in (3.46) is valid as the partial sums of all derivatives are uniformly convergent. To calculate the coefficients c_{k+N} , note that by (3.27) and the convolution theorem, we have

$$|u^{(l)}(\theta)| \leq 2 \sum_{k=1}^{\infty} k^l |\hat{L}_k| \leq \frac{2\pi\gamma}{N^\alpha} \sum_{k=1}^{\infty} k^l \sum_{j=1}^q |t_j| \left| \frac{e^{-2i\pi(u_j+l/2)kN^{-\alpha}} - e^{-2i\pi(u_j-l/2)kN^{-\alpha}}}{2\pi k N^{-\alpha}} \right| |\hat{\phi}(\epsilon_j k / (2\pi N^\alpha))| \quad (3.49)$$

$$\leq 2\gamma \sum_{j=1}^q \rho_j^{l-1} \sum_{k=1}^{\infty} (k/\rho_j)^{l-1} |t_j| |\hat{\phi}(k/(2\pi\rho_j))| \quad (3.50)$$

$$\sim 2\gamma \sum_{j=1}^q |t_j| \rho_j^l \int_0^{\infty} k^{l-1} |\hat{\phi}(k/(2\pi))| dk \quad (3.51)$$

where $\rho_j = N^\alpha/\epsilon_j \rightarrow \infty$. Hence there exists a constant $C > 0$ such that

$$|u^{(l)}(\theta)| \leq C \sum_{j=1}^q \rho_j^l |t_j| \leq C(N^\alpha/\epsilon^*)^l \sum_{j=1}^q |t_j| \quad (3.52)$$

uniformly in u_1, \dots, u_q . Inserting (3.52) into (3.48) yields the estimate

$$\left| e^{-u(\theta)} \left(\frac{d^p}{d\theta^p} e^{u(\theta)} \right) \right| \leq C \prod_{l=1}^p (\|t\| N^\alpha/\epsilon^*)^{lk_l} = C(\|t\| N^\alpha/\epsilon^*)^p \quad (3.53)$$

Inserting this into (3.47) gives the following bound on the Fourier coefficients of $c(\theta)$

$$|c_{k+N}| \leq C\gamma \left(\frac{\|t\| N^\alpha/\epsilon^*}{k+N} \right)^p \quad (3.54)$$

and the corresponding bound on the Hilbert-Schmidt norm

$$\begin{aligned} \sum_{k=1}^{\infty} k |c_{k+N}|^2 &\leq C\gamma \sum_{k=1}^{\infty} k \left(\frac{\|t\| N^\alpha/\epsilon^*}{k+N} \right)^{2p} \leq C \int_1^{\infty} k \left(\frac{\|t\| N^\alpha/\epsilon^*}{k+N} \right)^{2p} dk \\ &= O(N^2 (\|t\| N^\alpha/(\epsilon^* N))^{2p}). \end{aligned} \quad (3.55)$$

This completes the proof of the lemma. \square

The quantity (3.29) is close to a Riemann sum. We need good estimates on the error in the approximation, which is the purpose of the next lemma. This error is generically of order $N^{-\alpha}$ in the mesoscopic regime. More generally, for the case of diverging interval length $l = L(N) \rightarrow \infty$ the error becomes of order $\log(L(N))L(N)N^{-\alpha}$.

Proposition 3.9 (Macroscopic to mesoscopic). *Consider the quantity (3.29). We have the uniform approximation*

$$E_{\mathbf{u},\epsilon} = \frac{1}{(2\pi)^2} \int_0^{\infty} k |\hat{h}_{\mathbf{u},\epsilon}(k)|^2 dk + O(\log(L(N))L(N)N^{-\alpha}) \quad (3.56)$$

Proof. We use the fact that the error in a Riemann sum approximation is given by the step size (here $N^{-\alpha}$) multiplied by the total variation norm of the function in question. Hence we have to estimate the quantity

$$\mathcal{E} := N^{-\alpha} \int_0^{\infty} \left| \frac{d}{dk} \left(k \left| \sum_{j=1}^q t_j \frac{e^{-2\pi i k(u_j+l/2)} - e^{-2\pi i k(u_j-l/2)}}{2\pi k} \hat{\phi}(k\epsilon_j) \right|^2 \right) \right| dk \quad (3.57)$$

Using the identity

$$\begin{aligned} & \left| \sum_{j=1}^q t_j (e^{-2\pi i k(u_j + l/2)} - e^{-2\pi i k(u_j - l/2)}) \right|^2 \\ &= 8 \sin^2(2\pi k l/2) \sum_{j_1 \leq j_2} t_{j_1} t_{j_2} \hat{\phi}(k \epsilon_{j_1}) \hat{\phi}(-k \epsilon_{j_2}) \cos(2\pi k(u_{j_1} - u_{j_2})) \end{aligned} \quad (3.58)$$

and changing variables $k \rightarrow k/\epsilon^*$, we see that it is sufficient to bound the quantity

$$I_{\epsilon, u} := \int_0^\infty \left| \frac{d}{dk} \frac{\sin^2(\pi k l) \cos(2\pi k(u_{j_2} - u_{j_1}))}{k} \hat{\phi}(k \epsilon_{j_1}) \hat{\phi}(-k \epsilon_{j_2}) dk \right| \quad (3.59)$$

uniformly in u as $\epsilon \rightarrow 0$. Computing the derivative yields four terms $I = I_1 + I_2 + I_3 + I_4$ coming from differentiating $1/k$, the two trigonometric terms and the functions $\hat{\phi}$, respectively. The contribution coming from the derivative of $1/k$ is bounded by

$$I_1 \leq \int_0^\infty \frac{\sin^2(\pi k l)}{k^2} dk = \frac{\pi^2}{2} l \quad (3.60)$$

To compute the contribution coming from the trigonometric terms we change variables $k \rightarrow k/\epsilon^*$ so that the argument of the Fourier transform is diverging for $k > 1$. Then the contribution is dominated by the interval $k \in [0, 1]$ due to the rapid decay of $\hat{\phi}(k)$ and we get the bounds

$$\begin{aligned} I_2 &\leq C l \int_0^1 \left| \frac{\sin(2\pi k l/\epsilon^*)}{k} \right| dk = O(l \log(l/\epsilon^*)) \\ I_3 &\leq C 2\pi |u_{j_2} - u_{j_1}| \int_0^1 \frac{\sin^2(\pi k l)}{k} dk = O(|u_{j_2} - u_{j_1}| \log(l/\epsilon^*)) \end{aligned} \quad (3.61)$$

A similar estimate yields

$$I_4 \leq \max\{\epsilon_{j_1}, \epsilon_{j_2}\} \int_0^1 \frac{\sin^2(\pi k l/\epsilon^*)}{k} dk = O(\max\{\epsilon_{j_1}, \epsilon_{j_2}\} \log(l/\epsilon^*)) \quad (3.62)$$

Multiplying these estimates by the step size $N^{-\alpha}$ we get the error in the Riemann sum approximation

$$\mathcal{E} \leq C N^{-\alpha} \sum_{j_1 \leq j_2} |t_{j_2} t_{j_1}| l \log(l/\epsilon^*) \quad (3.63)$$

which completes the proof of the proposition. \square

A Appendix

We define the Fourier transform of any function $f \in L^1(\mathbb{R})$,

$$\mathcal{F}(f)(\kappa) = \hat{f}(\kappa) = \int_{\mathbb{R}} e^{-2\pi i \kappa x} f(x) dx. \quad (A.1)$$

The operator \mathcal{F} can be extended to a unitary transformation on $L^2(\mathbb{R})$ with the Plancherel formula:

$$\int_{\mathbb{R}} \hat{f}(\kappa) \overline{\hat{g}(\kappa)} d\kappa = \int_{\mathbb{R}} f(x) \overline{g(x)} dx,$$

for any function $f, g \in L^2(\mathbb{R} \rightarrow \mathbb{C})$.

We define the Hilbert transform of any function $f \in L^1(\mathbb{R})$,

$$\mathcal{H}(f)(x) = \frac{1}{\pi} \oint_{\mathbb{R}} \frac{f(u)}{x-u} du, \quad (\text{A.2})$$

where the integral is defined in the principal value sense. The Hilbert transform can also be extended to a bounded operator on $L^2(\mathbb{R})$ which satisfies:

$$\widehat{\mathcal{H}(f)}(\kappa) = -i \operatorname{sgn}(\kappa) \hat{f}(\kappa) \quad (\text{A.3})$$

where $\operatorname{sgn}(\cdot)$ is the sign function. In particular, the identity (A.3) implies that \mathcal{H} is invertible on $L^2(\mathbb{R})$ and $\mathcal{H}^{-1} = -\mathcal{H}$. Moreover, let us mention that if $f \in L^1 \cap L^2(\mathbb{R})$ is absolutely continuous (i.e. $f' \in L^1(\mathbb{R})$), then the Hilbert transform of the function f is differentiable on \mathbb{R} and

$$\frac{d\mathcal{H}(f)}{d\kappa} = \mathcal{H}(f'). \quad (\text{A.4})$$

We define the Cauchy transform of any function $f \in L^1(\mathbb{R})$,

$$\mathcal{C}(f)(z) = \frac{1}{2\pi i} \int \frac{f(x)}{x-z} dx. \quad (\text{A.5})$$

This function is analytic in both the lower and upper half planes, denoted \mathbb{C}_{\pm} . Moreover, its boundary values are given (in L^2 or pointwise if this limit makes sense) by the Plemelj-Sokhotski formula, for all $x \in \mathbb{R}$,

$$\mathcal{C}(f)_{\pm}(x) = \pm \frac{f(x)}{2} + \frac{i}{2} \mathcal{H}(f)(x). \quad (\text{A.6})$$

We define the Sobolev space

$$H^{1/2}(\mathbb{R}) = \left\{ f \in L^2(\mathbb{R} \rightarrow \mathbb{R}) : \int_0^\infty \kappa |\hat{f}(\kappa)|^2 d\kappa < \infty \right\}. \quad (\text{A.7})$$

This is a Hilbert space equipped with the inner-product

$$\langle f; g \rangle_{H^{1/2}} = \int_{\mathbb{R}} |\kappa| \hat{f}(\kappa) \hat{g}(-\kappa) d\kappa. \quad (\text{A.8})$$

There are other formula for the inner product (A.8) which do not involve the Fourier transform. For any functions $f, g \in C^1(\mathbb{R})$ such that $f, g \in L^2(\mathbb{R})$, we claim that

$$\langle f; g \rangle = \frac{1}{4\pi^2} \iint_{\mathbb{R}^2} \frac{f(x) - f(y)}{x-y} \frac{g(x) - g(y)}{x-y} dx dy \quad (\text{A.9})$$

$$= \frac{-1}{2\pi} \int_{\mathbb{R}} f'(u) \mathcal{H}(f)(u) du. \quad (\text{A.10})$$

It can be checked that formula (A.9) actually holds for any functions $f, g \in H^{1/2}(\mathbb{R})$, while (A.10) holds as long as $\mathcal{F}(f') \in L^2(\mathbb{R})$ and $\mathcal{F}(f')(\kappa) = 2\pi i \kappa \hat{f}(\kappa)$.

We define Ξ to be the Gaussian noise associated with the Hilbert space $H^{1/2}(\mathbb{R})$, see for instance [36]. That is $\Xi = \{\Xi(f)\}_{f \in H^{1/2}(\mathbb{R})}$ is a Gaussian process with covariance structure:

$$\mathbb{E} [\Xi(f) \Xi(g)] = \langle f; g \rangle_{H^{1/2}}. \quad (\text{A.11})$$

Observe that, if $\chi_u(x) = \pi \mathbb{1}_{|x-u| \leq \ell/2}$, then

$$\widehat{\chi_u}(\kappa) = e^{-2\pi i u \kappa} \frac{\sin(\pi \ell \kappa)}{\kappa}. \quad (\text{A.12})$$

So that, if we define $\widehat{Q}(\kappa) = \frac{\sin^2(\pi\ell\kappa)}{|\kappa|}$, it is easy to see that,

$$\langle \chi_u; \chi_v \rangle_{H^{1/2}} = Q(u - v) \quad (\text{A.13})$$

for all $u \neq v$. This proves that the log-correlated field Gaussian field G with zero mean and covariance function Q can be realized as $G(u) = \Xi(\chi_u)$. Similarly, we may consider a regularization of the field G . For any $\phi \in \mathcal{D}_0$, we let

$$G_\phi(u) = \int \phi(x - u) G(x) dx.$$

By Plancherel's formula, for any $\phi, \psi \in \mathcal{D}_0$ and $u, v \in \mathbb{R}$,

$$\begin{aligned} \mathbb{E}[G_\phi(u) G_\psi(v)] &= \iint_{\mathbb{R}^2} \phi(u - x) \psi(v - y) Q(x - y) dx dy \\ &= \int_{\mathbb{R}} e^{-2\pi i(u-v)\kappa} \widehat{\psi}(\kappa) \overline{\widehat{\phi}(\kappa)} \widehat{Q}(\kappa) d\kappa. \end{aligned} \quad (\text{A.14})$$

By definition of \widehat{Q} , we easily verify that the RHS of (A.14) is equal to $\mathbb{E}[\Xi(\chi_u * \phi) \Xi(\chi_v * \psi)]$, so that

$$G_\phi(u) = \Xi(\chi_u * \phi). \quad (\text{A.15})$$

References

- [1] L.-P. Arguin, D. Belius, and P. Bourgade. Maximum of the Characteristic Polynomial of Random Unitary Matrices. *Comm. Math. Phys.*, September 2016.
- [2] F. Bekerman and A. Lodhia. Mesoscopic central limit theorem for general β -ensembles. eprint = [arXiv:1605.05206](#), 2016.
- [3] I. Benjamini and O. Schramm. KPZ in one dimensional random geometry of multiplicative cascades. *Comm. Math. Phys.*, 289(2):653–662, 2009.
- [4] N. Berestycki. An elementary approach to Gaussian multiplicative chaos. eprint = [arXiv:1506.09113](#), 2015.
- [5] N. Berestycki. Introduction to the Gaussian Free Field and Liouville Quantum Gravity. <http://www.statslab.cam.ac.uk/~beresty/Articles/oxford.pdf>, 2016. [Online; accessed 2-Dec-2016].
- [6] T. Berggren and M. Duits. Mesoscopic fluctuations for the thinned Circular Unitary Ensemble. eprint = [arXiv:1611.00991](#), 2016.
- [7] A. Borodin. Determinantal point processes. In *The Oxford handbook of random matrix theory*, pages 231–249. Oxford Univ. Press, Oxford, 2011.
- [8] A. Borodin and A. Okounkov. A Fredholm determinant formula for Toeplitz determinants. *Integral Equations Operator Theory*, 37(4):386–396, 2000.
- [9] P. Bourgade, L. Erdos, H.-T. Yau, and J. Yin. Fixed energy universality for generalized Wigner matrices. *Comm. Pure and Appl. Math.*, 69(10):1815–1881, 2015.
- [10] P. Bourgade and J. Kuan. Strong Szegő asymptotics and zeros of the zeta-function. *Comm. Pure Appl. Math.*, 67(6):1028–1044, 2014.

- [11] J. Breuer and M. Duits. Universality of mesoscopic fluctuations for orthogonal polynomial ensembles. *Comm. Math. Phys.*, 342(2):491–531, 2016.
- [12] R. Chhaibi, J. Najnudel, and T. Madaule. On the maximum of the $C\beta E$ field. eprint = [arXiv:1607.00243](https://arxiv.org/abs/1607.00243), 2016.
- [13] T. Claeys and I. Krasovsky. Toeplitz determinants with merging singularities. *Duke Math. J.*, 164(15):2897–2987, 2015.
- [14] H. Cremers and D. Kadelka. On weak convergence of integral functionals of stochastic processes with applications to processes taking paths in L_p^E . *Stochastic Process. Appl.*, 21(2):305–317, 1986.
- [15] F. David, A. Kupiainen, R. Rhodes, and V. Vargas. Liouville quantum gravity on the Riemann sphere. *Comm. Math. Phys.*, 342(3):869–907, 2016.
- [16] P. Deift. Integrable operators. In *Differential operators and spectral theory*, volume 189 of *Amer. Math. Soc. Transl. Ser. 2*, pages 69–84. Amer. Math. Soc., Providence, RI, 1999.
- [17] P. Deift, A. Its, and I. Krasovsky. Toeplitz matrices and Toeplitz determinants under the impetus of the Ising model: some history and some recent results. *Comm. Pure Appl. Math.*, 66(9):1360–1438, 2013.
- [18] P. Deift, T. Kriecherbauer, K. T.-R. McLaughlin, S. Venakides, and X. Zhou. Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory. *Comm. Pure Appl. Math.*, 52(11):1335–1425, 1999.
- [19] P. Deift and X. Zhou. A steepest descent method for oscillatory Riemann-Hilbert problems. Asymptotics for the MKdV equation. *Ann. of Math. (2)*, 137(2):295–368, 1993.
- [20] B. Duplantier and S. Sheffield. Liouville quantum gravity and KPZ. *Invent. Math.*, 185(2):333–393, 2011.
- [21] L. Erdős and A. Knowles. The Altshuler-Shklovskii formulas for random band matrices II: The general case. *Ann. Henri Poincaré*, 16(3):709–799, 2015.
- [22] P. J. Forrester and S. O. Warnaar. The importance of the Selberg integral. *Bull. Amer. Math. Soc. (N.S.)*, 45(4):489–534, 2008.
- [23] Y. V. Fyodorov and J-P. Bouchaud. Freezing and extreme-value statistics in a random energy model with logarithmically correlated potential. *J. Phys. A*, 41(37):372001, 12, 2008.
- [24] Y. V. Fyodorov, G. A. Hiary, and J. P. Keating. Freezing transition, characteristic polynomials of random matrices, and the Riemann zeta function. *Phys. Rev. Lett.*, 108:170601, 2012.
- [25] Y. V. Fyodorov and J. P. Keating. Freezing transitions and extreme values: random matrix theory, and disordered landscapes. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 372(2007):20120503, 32, 2014.
- [26] Y. V. Fyodorov, B. A. Khoruzhenko, and N. J. Simm. Fractional Brownian motion with Hurst index $H = 0$ and the Gaussian Unitary Ensemble. *Ann. Probab.*, 44(4):2980–3031, 2016.
- [27] Y. V. Fyodorov and P. Le Doussal. Moments of the position of the maximum for GUE characteristic polynomials and for log-correlated Gaussian processes. *J. Stat. Phys.*, 164(1):190–240, 2016. With Appendix I by Alexei Borodin and Vadim Gorin.

- [28] Y. V. Fyodorov, P. Le Doussal, and A. Rosso. Statistical mechanics of logarithmic REM: duality, freezing and extreme value statistics of $1/f$ noises generated by Gaussian free fields. *J. Stat. Mech. Theory Exp.*, 2009(10):P10005, 32, 2009.
- [29] Y. V. Fyodorov and N. J. Simm. On the distribution of the maximum value of the characteristic polynomial of GUE random matrices. *Nonlinearity*, 29:2837–2855, 2016.
- [30] C. Garban. Quantum gravity and the KPZ formula [after Duplantier-Sheffield]. *Astérisque*, (352):Exp. No. 1052, ix, 315–354, 2013. Séminaire Bourbaki. Vol. 2011/2012. Exposés 1043–1058.
- [31] Y. He and A. Knowles. Mesoscopic eigenvalue statistics of Wigner matrices. eprint = [arXiv:1603.01499](#), 2016.
- [32] J. B. Hough, M. Krishnapur, Y. Peres, and B. Virág. Determinantal processes and independence. *Probab. Surv.*, 3:206–229, 2006.
- [33] X. Hu, J. Miller, and Y. Peres. Thick points of the Gaussian free field. *Ann. Probab.*, 38(2):896–926, 2010.
- [34] C. P. Hughes, J. P. Keating, and N. O’Connell. On the characteristic polynomial of a random unitary matrix. *Comm. Math. Phys.*, 220(2):429–451, 2001.
- [35] A. R. Its, A. G. Izergin, V. E. Korepin, and N. A. Slavnov. Differential equations for quantum correlation functions. In *Proceedings of the Conference on Yang-Baxter Equations, Conformal Invariance and Integrability in Statistical Mechanics and Field Theory*, volume 4, pages 1003–1037, 1990.
- [36] S. Janson. *Gaussian Hilbert spaces*, volume 129 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1997.
- [37] K. Johansson. Random matrices and determinantal processes. In *Mathematical statistical physics*, pages 1–55. Elsevier B. V., Amsterdam, 2006.
- [38] J-P. Kahane. Sur le chaos multiplicatif. *Ann. Sci. Math. Québec*, 9(2):105–150, 1985.
- [39] J-P. Kahane and J. Peyrière. Sur certaines martingales de Benoit Mandelbrot. *Advances in Math.*, 22(2):131–145, 1976.
- [40] O. Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.
- [41] I. Krasovsky. Aspects of Toeplitz determinants. In *Random walks, boundaries and spectra*, volume 64 of *Progr. Probab.*, pages 305–324. Birkhäuser/Springer Basel AG, Basel, 2011.
- [42] A. Kuijlaars. Riemann-Hilbert analysis for orthogonal polynomials. In *Orthogonal polynomials and special functions (Leuven, 2002)*, volume 1817 of *Lecture Notes in Math.*, pages 167–210. Springer, Berlin, 2003.
- [43] G. Lambert. Mesoscopic fluctuations for unitary invariant ensembles. eprint = [arXiv:1510.03641](#), 2016.
- [44] G. Lambert and E. Paquette. The law of large numbers for the maximum of almost Gaussian log-correlated fields coming from random matrices. eprint = [arXiv:1611.08885](#), 2016.
- [45] A. Lodhia and N. J. Simm. Mesoscopic linear statistics of Wigner matrices. eprint = [arXiv:1503.03533](#), 2015.

- [46] O. Macchi. The coincidence approach to stochastic point processes. *Advances in Appl. Probability*, 7:83–122, 1975.
- [47] B. B. Mandelbrot. Possible refinement of the log-normal hypothesis concerning the distribution of energy dissipation in intermittent turbulence. In M. Rosenblatt and C. Van Atta, editors, *Statistical Models and Turbulence*, volume 12, pages 333–351. Springer, New York, 1972.
- [48] B. B. Mandelbrot. Limit lognormal multifractal measures. In E. A. *et al* Gotsman, editor, *Frontiers of Physics: Landau Memorial Conference*, pages 309–340. Pergamon, New York, 1990.
- [49] J. Miller and S. Sheffield. Liouville quantum gravity and the Brownian map III: the conformal structure is determined. eprint = [arXiv:1608.05391](#), 2016.
- [50] D. Ostrovsky. Mellin transform of the limit lognormal distribution. *Comm. Math. Phys.*, 288(1):287–310, 2009.
- [51] D. Ostrovsky. Selberg integral as a meromorphic function. *Int. Math. Res. Not. IMRN*, 2013(17):3988–4028, 2013.
- [52] D. Ostrovsky. On Barnes Beta Distributions and Applications to the Maximum Distribution of the 2D Gaussian Free Field. *J. Stat. Phys.*, 164(6):1292–1317, 2016.
- [53] D. Ostrovsky. On Barnes beta distributions, Selberg integral and Riemann xi. *Forum Math.*, 28(1):1–23, 2016.
- [54] D. Ostrovsky. On Riemann zeroes, lognormal multiplicative chaos, and Selberg integral. *Non-linearity*, 29(2):426–464, 2016.
- [55] E. Paquette and O. Zeitouni. The maximum of the CUE field. eprint = [arXiv:1602.08875](#), 2016.
- [56] R. M. Pereira, C. Garban, and L. Chevillard. A dissipative random velocity field for fully developed fluid turbulence. *J. Fluid Mech.*, 794:369–408, 2016.
- [57] R. Rhodes and V. Vargas. KPZ formula for log-infinitely divisible multifractal random measures. *ESAIM Probab. Stat.*, 15:358–371, 2011.
- [58] R. Rhodes and V. Vargas. Gaussian multiplicative chaos and applications: a review. *Probab. Surv.*, 11:315–392, 2014.
- [59] R. Rhodes and V. Vargas. Lecture notes on Gaussian multiplicative chaos and Liouville Quantum Gravity. eprint = [arXiv:1602.07323](#), 2016.
- [60] R. Robert and V. Vargas. Gaussian multiplicative chaos revisited. *Ann. Probab.*, 38(2):605–631, 2010.
- [61] B. Rodgers. A central limit theorem for the zeroes of the zeta function. *Int. J. Number Theory*, 10(2):483–511, 2014.
- [62] E. Saksman and C. Webb. The Riemann zeta function and Gaussian multiplicative chaos: statistics on the critical line. eprint = [arXiv:1609.00027](#), 2016.
- [63] A. Shamov. On Gaussian multiplicative chaos. *J. Funct. Anal.*, 270(9):3224–3261, 2016.
- [64] B. Simon. *Orthogonal polynomials on the unit circle. Part 1*, volume 54 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2005. Classical theory.

- [65] B. Simon. *Trace ideals and their applications*, volume 120 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, second edition, 2005.
- [66] A. Soshnikov. The central limit theorem for local linear statistics in classical compact groups and related combinatorial identities. *Ann. Probab.*, 28(3):1353–1370, 2000.
- [67] A. Soshnikov. Determinantal random point fields. *Uspekhi Mat. Nauk*, 55(5(335)):107–160, 2000.
- [68] C. Webb. The characteristic polynomial of a random unitary matrix and Gaussian multiplicative chaos—the L^2 -phase. *Electron. J. Probab.*, 20:no. 104, 21, 2015.
- [69] C. Webb. Global spectral fluctuations in the Gaussian Unitary Ensemble. eprint = `arXiv:1510.05867`, 2015.